

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACCTC
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAAGCTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCTCGCCGCTGTG
CCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCCGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAAGCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAAGTCTTGCAGGTC
TAAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGCTGTGATAACTTCCTGACCTCCCTCCCAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCTCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAAGTCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAAGTGAACGCACTGTCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCTCAGGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTCCTTAAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGTTGGGATTCCAGC
CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAACCTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCCAGATGAAGAATTCATGGTTC
TGAAAATAAGGCTTCTGGTGTCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTACACACCACTGCCTCGTCTTTATTTACATTCTTAGAAACAA
CAATACAGGTAACCTCCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCTTCCGCAGAGGGGGCTCACAG
TGGCTACTGCAGCACTGTGAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAATA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAACTGGGAAAGATTGAGTGGAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACCTCCCTTG
CACCAGATCTTTGTAATAAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCTTGGCTTTTCAAGTCATTCCCCACGACAAGATGATGGTGGAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCACTCCTCAGC
AGTGGGGCGCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGTATAAGAAACC
TGGTCAAAACCCAACTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAATAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACCTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCACT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCATTTTATTCTCCAATGACAGGTTACATTTC
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAGCC
ACAGCCATTCGGCACTCCAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTG
GCCAGGCAGTTATTAGCGTGTCAAGTGTGGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACCTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGAA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACCTTAACTTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCCTCGTGTCTGTTTCCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGGCGTCTGTGTAAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

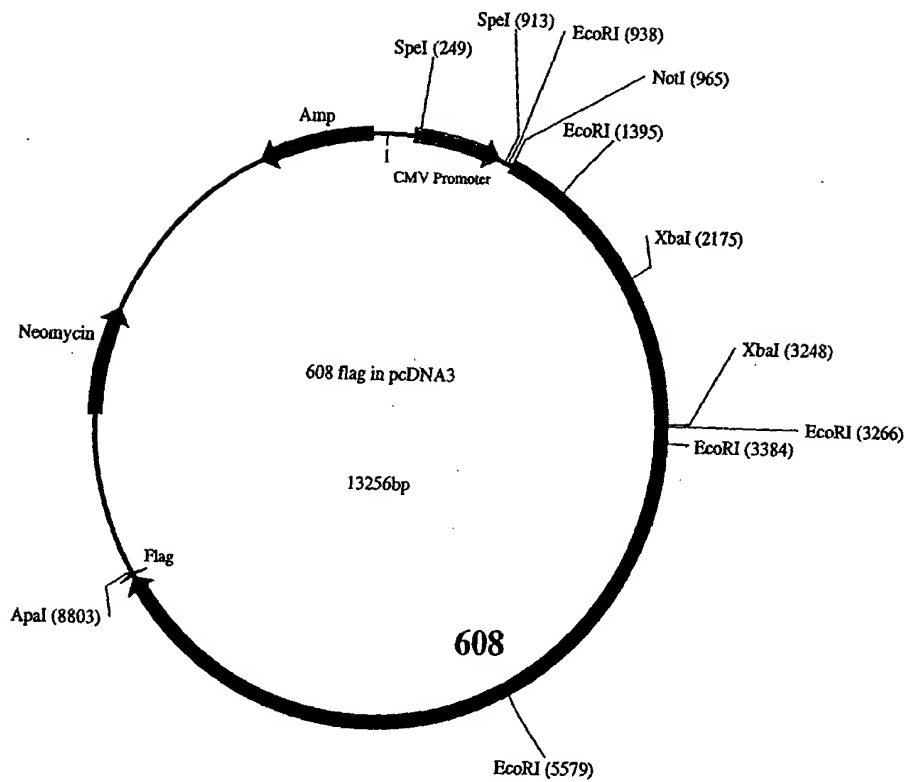


Figure 3

MQVRGREVSGLLISLTAVCLVVTGPSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
ADPFWFQQEKIVLQLNRTATTLSLQIQFSTDAQIALPRAEMRAERLKWTMILMMNNPK
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPLCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRLQVTPKDQGHYQCVAAANPSGADFSFVKV
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRK
NKHRDLIHRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
KPVPLAVPLVELTDEEKDASGMIPDDEEFMVLKTKASGVPGRSPTADSGPVNHGFMSTI
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
KNPHRTPVLRHRHRTVRPAIKGPANKNVSQVPATEYPMCHTCPSAEGTLVATAALS
VPSSSHSALPKTNNVGVI AEESTTVVKKPLLLFKDKQNV DIEITTTTKYSGGESNHVIPTE
ASMTSAPTSVSLGKSPVDNSGHL SMPGTIQTGKDSVETTP LPSPLSTPSIPTSTKFSKRKTP
LHQIFVNNQKKEGMLKNPYQFGLQKNPA AKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
TMAATQNGKTEVVSGARSLSAGKKQPFTNSSPVL PSTISKRSNTL NFLSTETPTVTSPTAT
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALT TADTPLAFSHSP
RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALET TLLSKSQESTTVKRASDTP
PPLLSSGAPPVPTSPPPFTKGVVTD SKVTS AFQMTSNRVVTIYESSRHNTDLQQPSAEAS
PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSWP EHKYQLKSYSETIEKGKRPA
VSMSPHLSLPEASTHASHWNTQKHA EKS VFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
GKAASFTVPANS DVFLPCEAVGDPLPIH WTRVSSGX EISQGTQKSRFHVLPNGTLSIQRV
SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
RPTVSWILANQTVVSETAKGSRKVWVTPDGT LIYNLSLYDRGFYKCVASNPSGQDSLL
VKIQVITAPPVIEQKRQAIVGVLGGS LKLPCTAKGTPQPSVHWVLYDGT ELKPLQLTHS
RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRV VILTVEEGETIPRIETASQKWTEVN
LGEKLLLNC SATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
YLCVARNKMGGDDLVMHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFN NVGMAEEGDYICSAQNTLGK
DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
NDRFTFHANRTL SIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
VIKATAIRH SKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

[illegible]

Figure 4

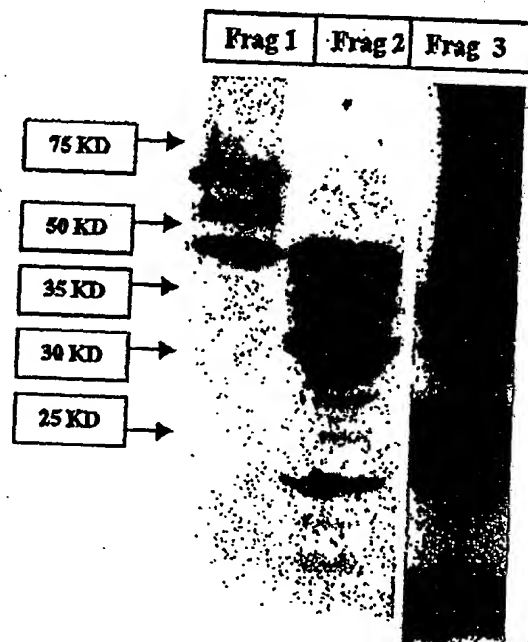


Figure 7

Nuc 1

↓
TTTGGAAACCAACCCAGATGCCCTCAACAGAGAAAATGGGCCAGAAAATGTGGTGCATTTA
TCCAATGGAATACTACTCAACTTATTAACCAACCACTTTCATAAAATTTTAGGCAAAAT
GNATGGTCTGNAGGATCTTGAGTGAGGTAACCAATCACAAAAGAACACTCATGGTATGC
ACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTIAAAAGGGAAGAAGACACATAGC
TTTTTGTGTGTATAATATTAAGATGGAAAATTTGCCAGTGCTGTTTGGCTTATGAGTGAATCT
TGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAACGTAGGAAGAAGTAGTTAATCA

Nuc 390

AGGTGACAAAAGTATCCTGACACATTGGGAAAAGACCACAGTCCAGGAAACTGAGTCTTAA
GGATTTCATATTAACCTCAGTTCCCATGTGCAGCTCTGAGACTTTGGCAGATCAGACACTT
AACTTCACCAGCTTCCTACACAGAGCAGTTACTATCCTTGCACCTTCACACATGGAGTGTGA
CCATTAAAGCTGCACGTGAAACATGAGTCTGACTTGTAAATAATCTTAAAAATACAAATTGTGT
TGTAAGATATGTGACCAAAAGAGCATGGTCACTGCTATTAACCTTTGATGTTCTATGGACTCT
TAATTTTATGGTAGAAAATGTCAACAAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAG
GATGGCCTTTTCAGTTTGTAAAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCA
ATTAAACCCCAAGTAGGCATAATTTAATGCTTACTTCATCAGAATATGTCTAATGAC
TCTTCTAAAAAGACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTACTGACATAAT
AAATAGGAQAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGCTCT
ATAGGTTATATAAGAGGTGGGAAAAGGAATATTGAAACATCTAGAAGTAAAATGATCCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTGTGGTTTGGCT
GGAGCTGAGGCTCTGGAAGGCTCGACCTTGGTTAGAAAATAACCTACCGAATACAGAGCTAT
GACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCCAGAACATACTA
GATGCCATATTTCCAAGGCAAGTGTCCACATGGGGACATCTTAAGAATATGGTTGTCTCTG
CAGTGCTAAGGACCTTGTGTGCTGCCACACAGGCTCTCCAGGTTAGTGCTAACCTCTGACTGC
TTGACTCTTTAATTTCTAOCCTGATCATTAATGACTAGAAAATCACTTGGTGTATTAGCAACTG
GATATGGAATATTACTAATTTGTACCCAAGCCAGGCCACCTCAGCTTTGGCAGCTCCATTC
ATTCTGTGGAGCCCACTGACGTTGGGTTTGAATCAACTGTACTGTTTCTACTTACAAGAGCG
ATTACCTGAGATGAGTCAATTTTCTTCACAAGTCTTTTGAAGAAGTCAATPAGACATATT
CTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATGTGTTCATGTATGCTCATGGA
TGCTATTATAATGTGGAAAATAAACTGACTTTAAAAAAAAGGCTTATGATACITGTTCACA
QAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAA
AAAGCTTTAGAAAATTTTGTAAACACAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTGA
TGGTGGGCTTGAGAGATGGAGCTTAAGAACTTACTTGTCTCCAGGAGCAGCTCTTCAGAAC
ATCTGCCTCAAAACATTTATCCCAATGCTCATCAAAAGGCTCACTCACATGTGCTTCAACC
ACAGGGATTAAACAGTCAATTTAGTTCACATTTCTCAAAAGGCTGGAAGGCTGTAGAGGAA
CAGGATGTATCAGGATAACATCCAACTTACAAAAGGATGTCTATAACCTCAACCAACA
AACAAACAACGACAAACAACCCATAAAATTTATCAAGGCAAAATGAACATAAGCCATATGCA
GAAAAAGTATTATATGTTCTCATTTGTGGGGTGTTTTCTTAATAGTCAAAATATGCAGAAT
ATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTTGTAGATTAGAGGACA
CAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAATGTATAATATGAGGACTGTA
TTAATAATTCTATTAAAGATACACAGCAAAAGAGTGTATCTTACTAACACACACACTTAC
ATAGAGAGAATAAAGTGATAGATACGTTTGTTTTATCTTCATGTAGCTGATAATTTCATAT
TGTACACCTCAAAACATAGATAACCAACAAGAGGAAAGAGGATAGGTGCCTCTCCAGGGC
GGAAGAGTACATTGGAAGTCAGACACCATTTGTGTAGATGTACCACATGGAGGAGCTAGA
GAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCCTATAGGTGGAACAACATTATGAGCT
AACCAGTACCCCGGAGCTCTGACTCTAGCTGCATATATATCAAAAAGATGGGCTAATCCG
CCATCACTGGAAGAGAGGCGCATTTGGACTTGCAACTTTATATGCCCAAGTACAGGGGA
ATACCAGGGGCCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGG
GGACTTTTGGTATAGCATTGGAATGTAAATGAGTTAAATAOCTAATAAAAAATGGAAAA

Nuc 2317

AAAAAAAAAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTT
GCAAAACAGAAAGGGGAGTCACAAACCCAGGACAAACACAGTGATTGAAGGCTCTTTGAA
TGTTATTGCTGTTGTTGTTACCAACCATCATTAGCATATATTCAATTGTGAAAACCTTACGGGGT
CTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTGACGATATTGGCACCACTACCA

[illegible]

GCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAAGGTTCCAGGAAAAAGGAAAGT
ACTGGAGGGGAGTTGGGACAAAAGCAGCGAOCAGGGGAACATCGCTTCAGTGACTGAAGCCA
GGCAAAAGGAAGCGGAAGGATTATATGTAGCTCTGGGAAGCTTTCATAAACACTGATGAAGTGT
TTGTGCAAGCAAGCAATTTGAGGAGCAAAACGCTGGGAAGCTCGGAAAAGGAGTGATCGATT
AGTACTTGTAAAGTTTAGGTGAGTTTGAGAACTAACTAACTATACTATTGAGGGAGAAAGCAAG
AGCATTCCAGCAGCAGCAGCAGCAGCAGCAATCAGATAAAGGAAAGCTTTGGTTAGTTTGGAA
ATGTATGATAACATTAAATAACAGAAAGCGCTCCAGTTCTCTGAAGAGTCACTCCCCAGCTA
GTGAAGACTAAGCTACTAAGGCTTTTGCTCCCGTTGGAAGCAAGAAAGTTCTTCAATCAGG
TGAAGGCTCTCCTCAGAAAGATTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGAC
AGAAGAGCTCAGGTATTGCCAACTCTTT

TGTTAAATACAGTTTGAAGGCTTAAGTGTAACGGGAACCTCATGTGGTATTCAATTAACGGCTCT
CTTCTCTTATAACTAAGCTTAAGGTGCATATAGTCTCTTCTGTTTCCAGCTACCTTGTAC
CATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATCAACGCAGAGATATT
CAAAAATATTTCAGTGATGTAACAGTAGCAGGTGTAGGCATAGAAGTAATCATTAGTAAATC
TTAATTGGGTGTAAGCTCATTCATACAAAGCTCCAGGTTGGGAGGGATCACTGAGCCCTCCG
CACGTGGCGGGTTAAAGATATTTTCTAAACAAGAGAAGCAGAATTCCTCTCTTGCCATGCTCC
CCATCACTGTGTCAAGTAAGCAGAGGGGTGTTCCTCAAGCAGAGAAAAGAGCAGACAGTGT
TGCCCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAAGTTTACTGTCTCTCCCGTCAIT
AGTTGGGCTGTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAA
GAAATTGACCTTTTCTTAGTGGGTTATTACAGCTGTAAAAGTATTTTGGAAAGGTTAAG
CCAAATTAATAAAACACATATTAAATAATCAATGTACAAAAATTGATCATATAAAGAA
TACATTTCAATAAATGCAATGTGAAAAAATATATAATTTTATTCTATTTACTGGTGCAAAG
TTTTCTAAAATTGCACATGTACTATTTTTATATTTATAAAAAATTTTTTAAATGTATATAAA
AGTGTAAAGGCTCTGGTCAAACAAGAGAGTTAAATTTACAAACTTTAATTGTCCCGAT
AACAATTATATGATCTCTAATGACAGGGATCTCTGCTTTTCATTGGGAAATGAGAAGCTATG
AAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCAAATGGGAAGCTAGCACAC
ACTGGTTTATAAAGAGAACAGTTTCTTGAGTCTATGCAAGTTTACCTCTAGGGAATAAG
AGTTTCTCTTTCTCCAGATTTTCACTAGCATTTGTTGTCATCATTTTACTCTTGATGATGAG
CATTATAAGTGGAATAAAGATAGGATCTCAAAGGAATGTCAATTTGGATGCCCTGAACAA
CTTTCAGGCTTTCTTTCAAGTTCACTAGTCTATTTCATTTATTGGATAATTGGGGGATGGTGT
TAATTTTTTGGCAGTCTTATGGAATTCAAAAAAACAAAAACAAACAAACAAACAA
ACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAAACAAAGAGAAATCTGCACA
GAATTTATGCTACATTGTTTCATTATTCAACGACAGCCAAAGAAATGTGGAACCAACTTACGTA
GCCGTCAAATAATGAAACGGATAAAGAAAAATGTGGAATGTGTACAACAGAGTCCATGTG
GCCCTAAAAGAGATGAAATCATGACATAATGCAAAAATGGTCAACCTGGAAATCAAATTTG
GGCTAATCAAAACAAGACAGAGCTCAAAAAGGAAACACCGTGTAGCTCTCTGACAAACA
GAAGCTAGATTACACTTGTACGTGGCATGTGTGTTTAGAATTTTATTTAGTTATACACT
ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTGGGG
TGGGTAGGAGAGAAAAGCAATGAGAGGAGTAAATAAGAACGAAGCATAGTAACATAGGTG
CCAGGATGAAATGCAATTTGTATGCTAATCAATCAACAGACAGGAGGCCACAGTTCA
AACCGAGGTGAAATCCAGCAGAGAGAAGGGGAAGTAGACACAAAGTTTGGCCACTAAC
CAAGAAGCCATTTCAGATTGCTGCTGCTGGTGGGAGGGGGTTCCAGTTTCTCCAGTCTGAC
ACTGTGTATAACAACAGTTTGACAAATACAAAGTTGGCATGATGGATGGTTTTTGTGCTATT
TTTCATTTTTTTCTTACTGTTTTGTTGTTGTTGGTGGTGTGTGTGTGTGTGGCTGTGGTTTT
ATTGTTTCTTTTGAGAGAGAGAAGGAAACATGAAATTGGGTTGGGTAGGAAGCTGGAAACG
ATCTGGGAAGAAGTTGGGGAAAGAGAAAAAATTGTATGGAGCATATTTAAACAAACAAACA
AACAAACAAAGGTTTCATTTTGCCAAAAAAGGTTGTAATTAATTAACAGGTTACGACT
CTTAAAGAAAAATATCCCAATTATTCOCAGAGTTGCTATGTATGCTGTGCCCTAGGACTTTG

CTTGAACTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAGAAGAGAGGCAGGGAA
GTCAGCTGCTTGTGATCTCCCTCACTGCATCTGCCTGGTGGTCACCCCTGGAGCAGGGTC
TGTCTCGCGGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTGGGGAOCTGAOC
TCATCCCAGACGGGCATCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

GGACCTTGGCTGATCTCCTTCTCAGAGAGGGGAACCACTGATTTTCTGGTACTTTGCCCCCC
AAACAACCTGTGATTACTTTTAATAGTTTTCTTCTAAAAATGGGTTTCATACAAACCTTATATTG
TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACCTTGAAGCCCTCTTAGTT
GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTTAAATGAAAAAGAT
ACAGTTTCTGTAAACAATCATTATGATACCAAGGAAGGAATTTGTCATTGGAATATTTTAA
AAAAGCACTTCTTTTGGCAATTTATAAAATACCCATTACAAAATGGCTTACTTAAAAATCTGTG
CCTTACTAAATCTGACAAATTATGGTGATATTTTGAAGGTTTATGAAAATTTGTTTATGTGT
ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT
ATGCCAAACACTATTGTTTCACGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT
TATAACCGTATACCCAGTGCCTTGGAAAGCGATTGCAAACAGTAAAGACTGACATTGTGTT
CTCCCTATGAGGGAGGGGGCCTTGGGCTGAGCACTTTGCAATGAGCATTIGCTCATTGTGCT
GGCAGGTTTTATGATAACTTGACCAAGTCAAGAGTCACTGGAGAGGAAGGAACTTCAACT
GAGAACATGGCTGAAGAAGATCAGATTATAGCAGGCCCTGTGGGCATTTTCTTAATTAG
TGATTTCATGGGGCAGGGCCAGTCCATTGTTCTGGTGATACCATTTTCTCAGGCACCTATAAAA
AAAAAAAAAACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGGCCAATGA
TCTCTGCATCAGCTCCTGCGCTCCAGGTTCTACCCATTTTGAGTTCTGTCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAAGTATAAGCCAAATAAAATCCTTTCTTCCCAACTTGCTG
TTGGTCATGATGTTTTCATCACAGTGATAATAGTCCCTCATGAAGATGCTGGTGTTTATAACA
CCTTTGGACTAAATCTGTATTCTATAGCTGAGGAAAAATGGAGCATAGAAAGTCTCCAGA
CTACACCAAGAGTGTAATCTGGGCTTAGCTTAGAATCACACCCACGTGCATCTCACTGCC
GGGGCTTCTTAAACCGGAACACAGTTGTAAAAGGGAATTTTCTGTTTGTTCATTTTGACA

TGTTGGAAGCTTTAAATTGACGATTCATCTGAAGCTGAAAATGATTTTTTTTCCAGGTATAACAGC
 CTCCTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATGCTGCACA
 GCAATGGCATTCAAGAGTCAAGTGAACAAGACCTTCTGGGGCTTGCAGTCTCTTCAG

GTGAGATAGGTAGAGGGTGTATGGAGGCTGAGAAGAGAGGTTGCAACTGTGGGTTATACCC
AAAAGCTGCTGATTCCCGTGGGAGACATTCTATAAGCATTCTATAAACTAGAGGCAGATA
TCAAGGAAGGATTTCAAATTGTATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATG
CTGGGAAAAATGCTTACACAATTGCGAGGCACTAATTATTAGGATCTCCAATAGCCACATAA
AAGACACAGCATGTGGGCGAGACACCTGCAATTCTGTCTCGAAGACACTGTTCAGAAAT
CCGACAGACTCATTTGGCCAAACACTCTATTCAATCAATGAAGTGCATATTCAGTGACAAA
ACTTGACTCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACT
CATGCATGAATAAGGGATCCAGAGAGAAAGGAAAGAAAAAGGAAGGAAAGGAAGGAAGG
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
GGAGAAGAAGAGAGAAGAAAGGAAAAATAATAAAATTTTCAGGGATTATTACACCTTTAA
TTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
AAGGGCAGTCTTTAACATAGTAGCATTTTATAAAACCATTACAAATTTTGAGTTTCTCTAC
TTTTATCTCTACCATCTTCAAACCTGAAACTACAATTATCCACAAATGAAGAAAAATGC
TGTAAGAGTTTTCACACAACGGAAGTGGGAAACTTAAGGATTAGACAAGTCTAACAAATGAG
AATGGGGAGAAACAAAAAGAGACTGCACAGGGAGGCCCTTTCTCTGCTTATAATCTTGACAC
TTGAGAAGCTAAATTGACGCTGCATGACTCACTCAATTTTAAGCAACAAATGCTGTGTGTC
ATGAAAAGCAATAAAGTACATATGTCCCAATAATTTCATCAAAATTTGCATGCCAGCAC
ATAATAGCAATCAAAGCAATAACACCCACTGTTCCACAGAGACTTTAAACATGAAACTGGA
ACTATGTCTAGTGTTTTGACTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTG
ATTAGGTCATCAGACACGCAATTTGGAACATGTATCTTCAGGAGGAATCATTCATGTATCT
GCATGAAATTTCCACACCTATGTTTATCTTCTTATGCCAGGTTTTTCTCTGATGGAGAAACAT
GGGTTTGAGGTTTTACTCCAGGTAACATTTAGGGAAGGAGCTGTCTATGTTCTCAGTTTGG
CTTTTATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAGAGATTACAAATTTA
GGTCAAAACAGAAAAATATGTAAGGTTATTGTTTTATTAGTATTCATGTTCTTTTCTT
TTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGAAAAATAATC
AGAGAGGGATTGATGAATGCTAAAGCATATCAATCAAAAAATTCAAAAATTTTGTCTTCTAATT
CAGAATCAATTAATTCATATTACTATAAAAAAGACAGCAGCCAGATGTGTGCCAGCTGAG
GAGTGGATAAACTGTGTAACTGTGAGTGTCTATGTAGAAACAGAAAAGGAGTGAAGGTTGA

TGTGGGCTGCAACATCTTGAAAAACATTGGGCTACATGATGGAAGCCAGGCCACAAAAAGCC
ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAAATACATGGATTCTTAGCAAAACAGA
GTAAGATGTTACTTAGGGTCAGGAAAAAGATTAAAAAATACTATTGATGTGGAATG
ATCTTAATTTGGGGAAAAAGACAATTCCTAAGACGAAATAGTTGAGGTAGATATAGTTAT
ATCCCTGTGGATATTGTAATAAACCCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGG
GCAGGAGGAAAGTGAAATGAGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTGCTG
GGTGGAAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATGTTTAGGGAAGGATC
AAAACCAACAGGAGTGAGGGAGACAATAGGAACCAATGAGAGGCCAAAGTTCATGGTCAA
TGTGTGTGGAGACACCATAATAAACTCCTTTTTGTTTGTCTAACTAAAAACCTAAAAATC
TAAAAACAAAAACATTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGGG
GAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTAACTGCTTTTACGTTT
TTTATTATTAAATTACTCTGTCTGTCATTATTATCATCATCATTATCGTCATCATCACTA
ATGCTTTTCAACCATACACAAATGTAGGCAGAAAGAGTGTATCCACTTAGTGAGGCAATCTT
GGAGAGGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAGGACAGTGAGAGGGAA
ATGAACAAGAAAAATGTGGACACATGCACAAAAATTCATAGTCCACTACATTACTTTG
TATTCTAATATTAAGAAAAATAATAAACCCATTCTGTGCACCTATCAOCCAGGCTCAACAG
TTATCTTTGGCCACAGATCCTGTCTCACTGTCATCCTGTCCACCTGAGTCCACTTAGCGTTCTG
AATCCAATCCAGGGCATGATGCTTACTCCTACACAGAACTAAAGATTAAAGAGAGTTTAA
AAGTAACCATGACATCTCTCTGTCTTTAGCGATAAGTTCCTAATATTTATGGCTGCTTGT
GTATGTTCTAATTTCTCTAATATTGTACATTAGTTGGCAACTACTTTGTTTGAATTGAGT
TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTTCTATATTATAAACTTTTCTCTCT

Exon 4 ? (Nuc 11286) ↓

TTGTGAAAGTTCTTTGAGAAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATG
AGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGG
TTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTACGGACTCACTTGC
TCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCATCCAGACACATTGTCTCTTT
GAGCTATCTCCAGATATTTAAAAACCTCCTTCATTAAGNAOCTGTACTTGTATGATAACTTCATTG
ACCTCCCTCCCAAAAGAAATGGTCTCTCTATGCCAAACCTAGAAAGCCTTTACTTGCATGGAA
ACCCATGGACCTGTGACTGCCATTTAAAGTGGTTGTCCGAGTGGATGCAGGGAAACCCAGGTA

Nuc 11680 ↑

ACTATCTGTGTTGTTGTTTCTTTTTTATARKACGTATTTTCTCAATTTTATTAGAAATGA
TATCCCAAAAGTCCCCATAAOCCTCCCCCCTTCCCTACCTACCCATTCCCATTTTTTGG
CCCTGGCATTCCCCTGTACTGGGGCATATAAAGTTTGGGTGTCCAATGGACCTCTCTTCC
AGTGATGGCCAACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG
TACTGGTTAGTTCATAATGTTGTTGACCTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

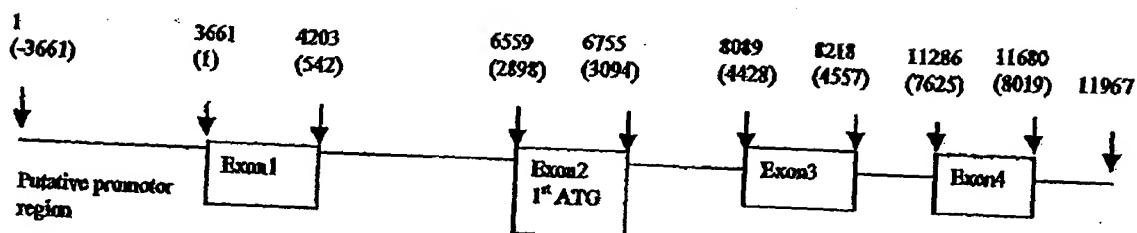


Figure 10

cDNA_rat		ogagagacgacagaaggttacggctgcgagaagaogacagaaggggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgtctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgactgaggccaggcaaaaogogcggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggaacctttcatagacactgatgacacgtttacgaaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaogcctgggccttcggaaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	ottgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagttccagcagcagcaggaaagccttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaaatggatgatagcattaaaaataacagaagcgcctccagg
genomic_hu	1	-----
cDNA_rat	401	tototgaagottoagtcctccocagctgaaagocagaaaagactaagoccaa
genomic_hu	1	-----
cDNA_rat	451	taagoccttttgatccctttggaagcaaagaactttccttcctgggggtga
genomic_hu	1	-----
cDNA_rat	501	agactctcctcagaagatttcctgtctctgcctatgtttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcaggtgagaggcaggggaagtcatg
genomic_hu	1	-----
cDNA_rat	601	oggttggttgatctcctcctcctcctcctcctcctcctcctcctcctcctc
genomic_hu	1	-----
cDNA_rat	651	gcaggggcctgtcctcctcctcctcctcctcctcctcctcctcctcctcctc
genomic_hu	1	-----
cDNA_rat	701	tgtacatttcggtaacctgacctccatccagatggcatcccgccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaatttaggatataacagccttactagattgacagaataacg
genomic_hu	1	-----
cDNA_rat	801	actttgatggcctgagcaaacctggagttactcatgotgcaagtaattggc
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcagtccttgcaagt
genomic_hu	1	-----
cDNA_rat	901	ottaaaaatgagctataacaaagtcocaaatcattoggaaggataactttct
genomic_hu	1	-----

cDNA_rat	951	acggactcgggagcttggtccggttgacactggatcacacaacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaaccctgaggccttttatggacttacctogctccgcttggtaca
genomic_hu	1	-----
cDNA_rat	1051	tttagaaggaaaccggtcacaaagctccatccagacacatttgtctcat
genomic_hu	1	-----
cDNA_rat	1101	taagotatctccagatatttaaaacctctttcattaagtacctgttcttg
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttctcgacctccotccocaaaagaatggtctctacatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacotagaaagcctgtatttgcattgaaacccatggacctgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	atttaaagtgggtgtgtgagtgatgcagggaacccagatataataaaa
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaagaacagaagotcttccagtcctcagcaatgtccctttgcat
genomic_hu	1	-----
cDNA_rat	1351	gaacccacaggatctctaaaggcagaccctttgctatggtacccatctggag
genomic_hu	1	-----
cDNA_rat	1401	ctttcctatgtacaaagccaaaccattgatccatcactgaagtoaaagagc
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu	1	-----
cDNA_rat	1501	catagaaaccccttggtcctgtctttgaacatgacanannttctggaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatggtctgtagtatccaaaagccatcaaggacatcacca
genomic_hu	1	-----
cDNA_rat	1601	actgoattcactgaagaaaatgactacatcatgctaaatgogtoattttc
genomic_hu	1	-----
cDNA_rat	1651	cacaaatottgtgtgcagtgtagattataatcacatccagccagtgtggc
genomic_hu	1	-----
cDNA_rat	1701	aaottctggctttatacagtgactctcctotgatactagaaaggaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagcttacccgagactccttcaactgtcttctagatataaacaggtggtot
genomic_hu	1	-----
cDNA_rat	1801	taggocctgaagacatttttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgcagotgaacagaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagcacattacagatccagttttccactgatgctcaaatcgc
genomic_hu	1	-----
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaattggaocatgatcc
genomic_hu	1	-----

cDNA_rat	2001	tgatgatgaacaatocaaaaotggaacgcactgtoctggttggcggcact
genomic_hu	1	-----
cDNA_rat	2051	attgccotgagctgtocaggcaaaaggogaoccttcacctcacttggaaatg
genomic_hu	1	-----
cDNA_rat	2101	gottotagotgatgggagtaaagtgaagagccoccttacgttagogaggatg
genomic_hu	1	-----
cDNA_rat	2151	ggcgaatcctaataagacaaaaatgggaagtgggaactgcagatggctgac
genomic_hu	1	-----
cDNA_rat	2201	agctttgatgcaggtotttaccactgcataagcaccaatgatgcagatgc
genomic_hu	1	-----
cDNA_rat	2251	ggatgttotcacatacaggataactgtggtagagccotatggagaaagca
genomic_hu	1	-----
cDNA_rat	2301	cacatgcagtgagggtocagcacacagtggttacgggtgagacgctcgao
genomic_hu	1	-----
cDNA_rat	2351	cttccatgcctttccacgggtgttccagatgottotattagotggattct
genomic_hu	1	-----
cDNA_rat	2401	tccagggaaactgtgttctctcagooatcaagagacaggcaaatcttta
genomic_hu	1	-----
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgcacaaagatcaagggtcat
genomic_hu	1	-----
cDNA_rat	2501	taccaatgtgtggotgccaaacooatcaggggcccacttttccagttttaa
genomic_hu	1	-----
cDNA_rat	2551	agtttcagttcaaaagaaaggccaaaggatggttgagcatgcacgggagg
genomic_hu	1	-----
cDNA_rat	2601	caggtggatctggacttggagaacocaaactccagtgttcccttaagcag
genomic_hu	1	-----
cDNA_rat	2651	coagcatctttgaaactctotgcacagctttgacagggtcagaggctgg
genomic_hu	1	-----
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaaacatagagacttaatac
genomic_hu	1	-----
cDNA_rat	2751	atcggcggcggtgggattccacgctccggcgattcaggagcataggagg
genomic_hu	1	-----
cDNA_rat	2801	cagctccctctctctgctcggagaattgaccgcgaacgctgggcagcact
genomic_hu	1	-----
cDNA_rat	2851	tctagaaaaagccaaaaagaattctgtgccaaaaagcaagaaaatacca
genomic_hu	1	-----
cDNA_rat	2901	cagtaaagccagtgccactggotgttccctcgtgggaactcactgaogag
genomic_hu	1	-----
cDNA_rat	2951	gaaaaggatgcctctggcatgattctccagatgaagaattcatgtttct
genomic_hu	1	-----

cDNA_rat	3001	gaaaactaaggcttctggtgtccaggaaggtcaccaactgctgactotg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatcatgggttttatgaogagtatagcttctggcacagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tceactgtgaatocacaaacactaocactctgagcaccttctgatttcaa
genomic_hu	1	-----
cDNA_rat	3151	attatttagtgtaacaaacgggtacagctgtgacaaagagtatgaacccat
genomic_hu	1	-----
cDNA_rat	3201	ccatgcaagcaaaatagaagatacaacaaacccaaacccaaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	ttcccatcagtagctgaaattcgagattctgtcaggcagggaagagcatc
genomic_hu	1	-----
cDNA_rat	3301	ttcccaagtgcaacacccgtgaacagggggaacatggctacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	atacacaacacatatagtagctttaccagcaagccagtagcagctctgcag
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaaacagaaagttatggacccagatccctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacotagcagtagtgacatctcttctcacactactgagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctccagtcaccccttcaggttcacacaccaactgcotogtctttattt
genomic_hu	1	-----
cDNA_rat	3551	cacattcctagaaacaacaatacaggttaacttcccttgcagggcaactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttggagcagagggagagttaaaacccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccagttctccagggcatagacacaggactgtgaggccagcaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggaacotgctaacaataatgtgagccaggtccagccacagagtacco
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgagaggggtcaagtggtactg
genomic_hu	1	-----
cDNA_rat	3801	cagcactgtcagttccaaagttccatccacagtgccctcccaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttgggtcatagcagaagagtctaccactgtggtcaagaaaccaot
genomic_hu	1	-----
cDNA_rat	3901	gttactatttaaggacaacaaatgtagatattggataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatattccggaggggaaagtaaccaogtgattoctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgaattctgtctcaaacatctgtatccotggggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgcctgggaccatccaaactgggaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaaacaacaccacttccagcccccctcagcacaccctcaataoca
genomic_hu	1	-----

cDNA_rat	5251	tgagaagggcaaaaggccagcagtaagcatgtcccccaacctoagccttc
genomic_hu	1	-----
cDNA_rat	5301	cagaggccagcactcatgcctcscactggaatacacagaagcatgcagaa
genomic_hu	1	-----
cDNA_rat	5351	aagagtgttttgataagaaacctgggtcaaaacocaaottccaacatct
genomic_hu	1	-----
cDNA_rat	5401	gccttacgtctctctacctaagactotattgaaaaagccaagaataattg
genomic_hu	1	-----
cDNA_rat	5451	gaggaaagggtgcaagctttacagttccagctaattcagaogtttttctt
genomic_hu	1	-----
cDNA_rat	5501	ccttgtagaggotgttgagagccactgcccatcatccactggacooagagt
genomic_hu	1	-----
cDNA_rat	5551	ttcatoagganttgaaatatccaaagggacacagaaaagcoggttccacg
genomic_hu	1	-----
cDNA_rat	5601	tgottoccaatggcaccttgtccatccagagggtcagttatccaggaccgt
genomic_hu	1	-----
cDNA_rat	5651	ggacagtacctgtgtctgtcatttaatcocaotgggcgtagaccattttca
genomic_hu	1	-----
cDNA_rat	5701	tgtototttgtctgtgtgttttttacocggcaaggattttggacagacatg
genomic_hu	1	-----
cDNA_rat	5751	tcaaggagattcacagttcactttggaagtactgtggaactaaagtgcaga
genomic_hu	1	-----
cDNA_rat	5801	gtggagggtatgccgaggcctacggtttcctggatacttgcaaaccaaac
genomic_hu	1	-----
cDNA_rat	5851	ggtggtctcagaaacggccaagggaagcagaaaggtctggttaacacctg
genomic_hu	1	-----
cDNA_rat	5901	atggaacattgatcatotataatctgagtcctttatgatcgtggtttttac
genomic_hu	1	-----
cDNA_rat	5951	aagtggtggtggccagcaaccatctggccaggattcactgttggttaagat
genomic_hu	1	-----
cDNA_rat	6001	acaagtcattcacagctcccoctgtcattatagagcaaaagaggcaagcca
genomic_hu	1	-----
cDNA_rat	6051	tcgttggggttttaggtggaagtttgaaaotgcootgcactgcaaaaagga
genomic_hu	1	-----
cDNA_rat	6101	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu	1	-----
cDNA_rat	6151	accattgcagttgactcattccagatttttctgtatccaaatggaactc
genomic_hu	1	-----
cDNA_rat	6201	tgtatataagaagatogotccttcagtgaggggcacttatgagtgcat
genomic_hu	1	-----
cDNA_rat	6251	gccaccagctcctcagggtcagagagaagggtagtgattcttactgtgga
genomic_hu	1	-----

rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
o_hu	890	CtatGGAAGCAGAATCACAGTCCATAAAATGGAACCTTGGAsattagga
rat	7401	acatccGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGttcggagcgag
o_hu	940	atgtgaGGCTTTCAGATTACAGCCGACTTATCTGTGTGGcccgaaatgaa
rat	7451	ggaGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCCTAGAAATGCTGAG
o_hu	990	ggtGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAG
rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAAagtcatecgcccaagctg
o_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgccagctgg
rat	7551	gcaagcccgTAGCACTGAAGTCTCTGTGGATGGGAACCCCCACCTGAA
o_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTGGATGGTAACCCACACCTGAA
rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAcaaa
o_hu	1140	ataatCTGGATTTTACCAAATGGCACAGATTTTCCAATGGACCAcaag
rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAgcaa
o_hu	1190	ttatcagTATCTGATAGCAAGCAATGGTTCTTTTATCATTCTAAAcaaa
rat	7701	ctcggaaacagtcAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTGGC
o_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAGTTGGC
rat	7751	TAcategagaaactaatootgttagagATTGGGCAGAAGCCAGTCATTCT
o_hu	1290	TAtattgagaaattagtoaatattagaaATTGGCCAGAAGCCAGTTATTCT
rat	7801	GACATACGAACCCAGGGatggtgaagagcgTCAGTGGGGAACCGTTATCAC
o_hu	1340	TACCTATGCACCCAGGGAacagtaaaaggoatCAGTGGAGAATCTCTATCAC
rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
o_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAAATGGACTATG
rat	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAct
o_hu	1440	CCAAGTGGttATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATatt
rat	7951	GCATGAAAATGGCACGCTGGTCATCAAAGCAACAACAGCtcaagcaaaAG
o_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
rat	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTGGCCAggcagttattago
o_hu	1540	GAACTATATCTGTAAGGCTCAAAATAGTGTGGTCatacaactgattact
rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCCTCCCCGAATCATAAactacotACC
o_hu	1590	gttcCAGTAATGATTGTAGCCTACCCCTCCCCGAATTACAAAtogtccACC
rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
o_hu	1640	CAGGAGTATTGTCACCAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
rat	8151	CCTTGGGAATCCCCAAGCCAaAAGTCACCTGGGAGACGCCAAGACACTCC
o_hu	1690	CCTTGGGAGTTCCCAAGCCagAAATCACATGGGAGATGCCTGACCACTCC
rat	8201	CTGCTCTCAaaagcaacagcaagaaaaccCATAGAAGTGAGATGCTTCA
o_hu	1740	CTTCTCTCAAcggcaagtaagagaggacaCATGGAAGTGAGCAGCTTCA
rat	8251	CCCACAAGGTACGCTgGTCATTTCAGAATCTCCAAACCTCGGATTCCGGag
o_hu	1790	CTTACAAGGTACCCTaGTCATTTCAGAATCCCCAAACCTCCGATTCTGGga
rat	8301	tcTATAAGTGCAGAGCTCAGAACTACTTGGgactGATTACGCAACAACT
o_hu	1840	taTACAAATGCACAGCAAGAACCACCTTGGtagTATTATGCAGCAACG
rat	8351	TACATCCAGGTACTCTGACAGGAagggggagactaaaattocacagaagt
o_hu	1890	TATATTCAAGTAATCTGACATGAaataataaagtcacaacatctgggca

Figure 12

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGACAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTTCAGAAAGATACCTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAATATTGAGTTTATAAACCAGAGGTTTTTATGGGCTCAACTTTCTCCGCTGGTGCCTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGCCAGATGTAATAAAATGCAAAAAGATAGAAGTCCCTCTAGTGTCTCAGCAGTGT
CCACTTTTGATGAACCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAG
TGTGCCAAGCCAACCATTTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCT
GCTTTCATCTCTCCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCT
GGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCCATTGCATTCACTGAA
GAAAATGACTACATCGTGCTAAATACCTTCATTTTCAACATTTTGGTGTGCAACATAGATTACGGTCAC
ATTGAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTG
CTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTACC
AACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAGCTGAAC
AGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGA
GCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACAT
ACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCACCCACAGTGGAT
TGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCCTATGTCAAGTGGATGGACGGATCTTAATAGAC
AAAAGTGGAAAATTGGAACCTCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATAAGCAGC
AATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTGAAGCCTATCAG
GAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACTGGTATC
CCAGATGCCCTCTATTAGCTGGGTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAA
GTTCTAAACAATGGCACATTAAGAATATTACAGGTACCCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTGATTTTTTGGATTTTCCAAGTTTCAAGTCAAGATGAAAGGACAAAGGCC
TTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCATCCTATTGCTCATCTTAAGGAG
CCACCAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACA
AGTAAGAGGCACAACATATCGGGAATTAACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGG
GAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCCTGTTGGAG
AAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCAACCCAGTGGTCACC
CAACTCCCAAACATACCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATTTATG
GTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
AGTCCTATGACAAACATAAATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCT
GAAGAACCCACAGATTTCAAACGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATAAACCCA
ACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCATCCACTGTCTTTCCACTGCTACTTGGGA
GCAACTGAATTTCAAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCAATA
ACAGTAAGGACTATGATCAAAGATGTCAATGTCAAATGCTTAGTAGCACCACCAACAACTATTATTA
GAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCAC
TTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGCTCAT
TCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTTCAGACGCTTTGGGAGGCAGAGG
AAAATTGGCGGAAGGGGGGGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGC
ATTTTCAGGTCAACAACCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAAT
GTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCTTAGTC
CAGAATCCACTATTACTACTTGAGAACAAACCCAGTGATAGAGAAAACAACCCACAATAAAATATTTTC
AGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATACCCATG
GAAAAAATCACAAGTAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCA
GTGATTACATCGTCACTTTCAAGTGCTATCACAAGCCACCAATGACTATTATAGCCATTACAAGGTTT
TCAAGAAGGAAAATTTCCCTGGCAACAGAATTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGAAT
CAACATAAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCTTTACCACAG
AGACAAAAGTTCCCTTTTCATTTCAACCACTTTCAACAAGTGTGATGCAAATTCATCTAATACCTTG
ACTACCGCTCACCACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCC
TTCCACCCCTTAACCTATGCTTCCCTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATCATATCA
ACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTCACTTATGAAACCCAAACA
GAGAGATCTAGAGCACAACAATAACAAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGACCCAAAC
ATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACT

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

[illegible]

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGCCTTACTAGATTGACAGAAAACGACTTTGATGGCTGAGCAAACTGGAGTTACTCATG
AGCTTGGTTAGATTGATGGAACAGATTTTCTGGCTGACCAAACTGGAGTTACTCATG
AGCCTCACTAGATTGACAGAAAATGACTTTCTGGCTGAGCAGACTGGAGTTACTCATG
*** * ***** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGCACAGTAATGGCATTCAAGAGTCAGTGACAAGACCTTCTCGGGCTTCAGTCCCTTG
CTTCACAGCAATGGCATTCAAGCAATCCCTGACAGACCTTCTCAGATTTCAGGCTTG
CTGCACAGCAATGGCATTCAAGAGTCAGTGACAAGACCTTCTCGGGCTTCAGTCCCTTG
** ***** ** ***** * *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCAITCGAAGGATACTTTCTACGGA
CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCAITCGAAGGATACTTTCTACGGA
CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCAITCGAAGGATACTTTCTACGGA
***** ** * ***** ** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCGGGAGCTTGGTCCGGTTGCACTGGATCACAACCAATTCATTCATCAACCTGAG
CTCAGGAGCTTGACAGGATTGACATGGACACAACCAATTCATTCATCAACCTGAG
CTCAGGAGCTTGACAGGATTGACATGGACACAACCAATTCATTCATCAACCTGAG
*** ***** ** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAACCGGCTCACAAAG
GTTTTTATGGGCTCAACTTCTCCGCTGGTGCACITGGAAGGAATCAGCTCACTAAG
GCTTTTACGGACTCACTTGTCTCCGCTGGTACATTTAGAAGGAACCGGCTCACAAAG
* ***** ** * ***** ** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCCATCCAGACACATTTGTCTCTTAAGCTATCTCCAGATATTTAAAACTCTTTTCAAT
CTCCATCCAGACACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAACTCTTTTCAAT
CTCCATCCAGACACATTTGTCTCTTTGAGCTATCTCCAGATATTTAAAACTCTTTTCAAT
***** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGTACCTGTTCTGTCTGATACTTCT-GACCTCCCTCCCAAGAAATGGTCTCTTA
AAGTACCTGTTCTGTCTGATACTTCT-GACCTCCCTCCCTCAAGAGATGGTCTCTTA
AAGTACCTGTTCTGTCTGATACTTCTGACCTCCCTCCCAAGAAATGGTCTCTTA
*** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CATGCAAACTAGAAAGCCTGTATTTGATGGAACCATGGAACCTGTACTGCCATTT
TATGCTGACCTAGACAGCCTTTAAGCTGATGGAACCATGGAACCTGTACTGCCATTT
TATGCAAACTAGAAAGCCTTACTTTCATGGAACCATGGAACCTGTACTGCCATTT
***** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAGTGGTTGTCTGAGTGGATGCAAGGGAACCATGATATAATAAAATGCAAGAAAGACAG
AAAGTGGTTGTCTGAGTGGATGCAAGGGAACCATGATATAATAAAATGCAAGAAAGACAG
AAAGTGGTTGTCTGAGTGGATGCAAGGGAACCATGATATAATAAAATGCAAGAAAGACAG
***** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGCTCTTCCAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAGGCA
AAGTCCCTGATGCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAGGCA
TTCTTTTATARKAGTATTTCTCAATTTCAATTTAGATGATATCCCAAGATC-
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACCTTTGCTATGTTACCATCTGGAGCTTCTATGTAAGCAACCAATTTGATCCAT
AGCCGTTAGCTATGTTCTCAGCTGCACTTCTCAGTGTGCCAAGCCCAATTTGATCCAT
-CCCATAACTCCCTCCCTCC-CTTCCCTACCTACCATTC-CAATTTTGGC
** * * * * ** * * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CACTGAAGTCAAGAGCCTGTTTACTCAGGAGGCAATGGATCTGCTCCACCTCACTC
CCTGAAATCAAGAGCCTGATCTATCTGGAAGACAGTATCTGCTTCTATCTCTCCCT
CCTGGCATTCCTC-
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGATTTATAGAAACCTTTGGCTCTTGTCTTTGAACATGACNNNTNTCTGGAATA
AAGTTTATAGGACCTTTGGCTCTTGTCTTTGAATATGACAGATCTGCTGGAATA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGGCCGACATGGTCTGTAGTATCCAAAGCATCAAGGACATCAACCACTGCATTCCTG
AAGCTAACATGGTCTGAGTATCCAAAGCATCAAGGACATCAACCACTGCATTCCTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAGAAATGACTACATCATGCTAAATGOSTCATTPTCCACAAATCTTGTGTGCAGTGTAG
AAGAAATGACTACATCCTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATTATAATCACATCCAGCCAGTGTGGCACTTCTGGCTTTATACAGTGACTCTCTCTGA
ATTACGGTCACTTCAGCCAGTGTGGCAATTTTGGCTTTGTACAGTGAATCTCTCTGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TACTAGAAAGGAAGCCCGAGCTTACCGAGACTCCTTCAGTGTCTTCTAGATATAACAGG
TACTAGAAAGGAAGCCACTTGTCTAGTGAACACCGCAGCTCTATTACAATATAACAGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGCTCTTAGGCTGAAGACATTTTACCAGCATAGAGGCTGATGTGAGAGCAGACCOCTT
TGGCTCTTAGGCTGAAGACATTTTACCAGCATAGAGGCTGATGTGAGAGCAGACCOCTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTGGTTCCACAGAAAAATTTGTCTTGAGCTGAACAGAACTGCCACCACTTAGCA
CTTGGTTAATGCAAGACCAATTTCTTGAGCTGAACAGAACTGCCACCACTTAGCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CATTACAGATCCAGTTTTCAGTGTGCTCAATTCGCTTTACCAAGGGCCGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGTGCTCAATTCAGTTTACCAAGGAGCAGAGATGAGGC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGAGAGAGCTCAAATGGACATGATCTGTGATGAACATCCCAACTGGAAACGCACTG
CAGTGAACACAAATGGACTATGATTTCAAGGGATAACATCTAAGCTGGAACTACTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCTTGGTTGGCGCACTATTGCTGTGAGCTGTCCAGGCAAGGGAGACCOCTTCACTCTCT
TCTTGGTTAGGTGAACCGTTGGCTGAGTGTCCAGGCAAGGGAGACCOCTTCACTCTCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTACGTTAGCGAGGATGGGC
TGGATTGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTATGTCAGTGAAGATGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GAATCCTATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGATGCAG
GGATCCTATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGATAGTTTGTACACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTCTTTACCACTGCATAAGCAACATGATGCAGATGGGATGTTCTCACATACAGGATAA
GGGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACAGTGGTTA
CTGTGGTAGAAGCTTTGGTGGAGCCCTATCAGGAAATGGGATTCATCAGCAGTCTTCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGGGTGAGAGGCTCGACCTTCCATGCTTTCCAGGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAACACCTTGATCTTCCATGCTTCTACTGGTATCCAGATGCTCTATTAGCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGATTCTTCCAGGGAACACTGTGTTCTCTCAGCATCAAGAGACAGGCAATTTCTAACA
GGGTATTCCAGGAACCAATGTGCTCTATCAGTCATCAGAGACAGAAAGTTCTAACA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATGGGACCTTAAGATATTACAGGTACGCCAAAGATCAAGGTCATTACCAATGTGTGG
ATGGCACATTAAATATTACAGGTACCCCCAAGACCAAGGTTATTATOGCTGTGTGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGCCAACCCATCAGGGGCGGACTTTTCCAGTTTAAAGTTTCAGTTCAAAGAAAGGCC
CAGCCAACCCATCAGGGGTTGATTTTGTATTTCCAGTTTCAGTCAAGATGAAGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAAGGATGGTTGAGCATGACAGGGAGGCCAGGTGGATCTGGACTTGAGAACCCAACTCCA
AAAGGCCCTTGGAGCATGATGGAGAACAGAGGGATCTGGACTTGATGATCCAACTCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGTGTTCCCTTAAGCAGCCAGCACTTTGAACTCTCTGCATCAGCTTTGACAGGGTCAG
TTGCTCATCTTAAGGAGCCACAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGGCTGGAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC
AGGTTGGAAACACACCTCAAGCACAGTAAGAGGCACAACTATCGGGATTAACACTCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGCGGCGTGGGGATTCCAGCGCTCGGGCGATTCAAGGAGCATAGGAGGCAGCTCCCTCTCT
AGCGAAGTGGAGATTCAACACATGACGTTTAGGGGAGATAGGAGGCATTTCCTCCCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGCTCGGAGAAATTGAACCGCAACGCTGGGCGCACTTCTAGAAAAGCCAAAAGAAATT
CTGCTAGGAGAAATTGAACCCACAACTTGGGCGCACTGTTGGAGAAAGCTAAAAAGAAATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGCCAAAAGCAAGAAATATCCACAGTAAGGCCAGTGCCACTGGCTGTTCCCTCTG
CTATGCCAGACAAGCCAGAAATATCCACAGTGAGCCCAACCCAGTGGTCAACCACTCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGAACTCACTGACGAGGAAAGGATGCTCTGGCATGATTCTCCAGATGAAGAATTCA
CAACATACCTGGTGAAGAGACGATTCTCAGGCATGCTGCTCTACATGAGGAATTTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGTTCTGAAACTAAGGCTTCTGGTGTCCAGGAAGGTCAACCACTGCTGACTCTGGAC
TGGTCCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGACTGCTGACTCCAGAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CAGTAATCATGGTTTTATGACGAGTATAGCTTCTGGACAGAGTCTCACTGTGAATC
CAATATCTGATAGTCTATGACAAACATAAATTATGGCACAGAACTCTCCGTTGTGAATT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CACAAACACTACATCTGAGCACTTCTGATTTCAAAATATTTAGTGAACAAAGGTA
CACAAATFACTACCACTGAAGAACCAAGATTTCAAACTGTCTACTGCTATTAAACTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CAGCTGTACAAAGAGTATGAACCATCCATAGCAAGCAAAATAGAAGATACAAACCAAC
CAGCCATGTCAAAGATATTAACCAACCATGTCAAGCAAAATACAAGGCACAAACCAATC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAAACCAATCATTAATCTTTCCATC-----AGTAGCTGAATTCGAGATTCTG-CT
AACATTCTACACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTAGGACTCTGACA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGCA---GGAAGAGCATCTTCCAAAGTGCACACCTGTAAACAGGGGAAACATGGCT
GAGGGAGAGGGAAGAGGACATTTCC--AGTA-ACCCCAATAACAGTAAGGACTATGATC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACCTATGGCCTAOCACACATA--TASTAGCTTTACCAGCAAAGCAGTACAGTCTTGC
AAGATGNTCAATGTCAAANATGCTTAGTAGCAACCAACAAA-CTATTA-----TTAG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGCCATAAATCCACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCAGCAGAC
AGTCAGTAAATACCAACAATAGTCAT-----CAGACATCTGTAGAGTAAGTGAAGC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCTAGCTTCTCCAGTCACCTT
CCAGGCACAATCACTTCTATTCTCACACTACTCAATACTTAGCACCTCCAGTTCCCTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGTTCCACACACCACTGCTCTTTATTTCACATTCTAGAAACAACATACAGGTA
CAGATCCACACACAGCTGCTCATTTCTCAGTTTCCGATCCCTAGANNNAATAGTACAGTTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACTTCCCTTGTCTCAGGCACTTGGGAAGAGAGAGGACAAITTTGGAGCAGAGGGAGAGTTA
ACATCCCGCTGTTCAGACGCTTGGGAGGCAGAGGAAAATTGGCGGAGGGGGGAGATT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAACCCACATAGAAACCCAGTTCTCCGACGGCATAGACACAGGACTGTGAGGCCAGCAA
TCAGCCCATATAGAACTCCAGTTCTGCGAAGGCATAGATACAGCATTTTCAGGTCAACAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCTAGGGACCTGCTAACAAAAATGTAGCCAAAGTTCCAGCCACAGAGTACCTGGGATGT
CCAGAGGTTCTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGTCAATGTGACAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCCACACATGTCTTCCGAGAGGGGCTCAGTGGCTACTGCAGCACTGTCAAGTTCCAA
GTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTTGTCTTTCCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTTCATCCCAAGTGGCTCCCAAACTAAATATGTTGGGGTCATAGCAGAAGAGTCTA
GTGCTGTCTCCATCACTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAGAGTCTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCACTGTGGTCAAGAAACACTGTTACTATTAAAGGACAAACAAATGTAGATATTGAGA
CAACTCTAGTCCAGATCCACTATTACTTCTGAGAACAAACCCAGTGTAGANNNGAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TAATAACAACCACTACAAATATTCCGGAGGGGAAGTAACCAAGTGAATCTTACGGHAG
NNACAACACCCACAATAAATATTCAGGACTNGAAATTTCCCAAGTGACTCCAACCTGGTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAAGCATGACTTCTGCTCAACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTG
CAGTCATGACATATGCTCAACATCCATACCCATGGGAAAACTCACAAGTAACGCCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTCACCTGAGCTGCTGGGACCATCCAACTGGGAAAGATTCAAGTGGAAACACACCAC
GTTACCCACGTGTGTCTAGCAACCAATGAAGCTAAAAGAGATTCAAGTGATTACATCTGTCAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTCCAGCCCCCTCAGCACACCTCAATACCA-----CAAGCACAAAATTCCTAAAGA
TTTCAGGTGCTATCAOCCAGCCACCAATGCTATTATAGOCATTACAGGTTTTCAGAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGAAACTTCCTTGCCACAGATCTTTGTAATAACAGAGAGAGGGGATGTTAAGA
GGAAATTCCTGGCACAGACTTTGTAAATAOCCATACCCAAAAGGCAGATTAAAGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATCCATATCAATTGGTTACAAAAGAACCCAGCCGCAAGGTTCCCAAAATAGCTCTC
ATCAACATAAAGTTAGTTTACAAAAGCCACAGGTGTGATGCTTCTTAAACATCTCTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTACCCACAG-GTCAGAGTTCCOCTCAGATTCTACAACCTCTTGCACAGTCCSOCA
CTTTACC-ACAGAGACAAAGTTCCOCTTTCCATTTCACCCACATTTCAACAGTGTGATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCAGCTCTGTCTACAACATGGCTGCCACTCAGAACAGGGCACTGAAGTAGTATCAGGT
CAAATTCCTCTAATAOCTTGACTACCGCTCACCACACTACGACAAA--ACACACAA-T

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GCCAGAAGTCTCTCAGCAGGGAAGAACGCCCTTCAOCCAACTCTCTCCAG-TGCTTCC
CCTGGAAAGTCTTCCACAAAGAGGAGCTTCCCTTC-CCACCCCTTACCCCTATGCTTCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TAGCOCATAAGCAAGAGATCTAATACATTAACTTCTTGTCAACGGAAACCCACAGT
TAGTATTATAAGCAAGACTCAAGTACAAAAGCATCATATCAACGCCAACAGCAACCGC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GACAGTCTACTGCTACTGCTATGTCTGTAACCCACAGACAGATCCCA
AACAACTCTACCTTCCCTGCTATGTCTACTTATGAAACCCAAACAGAGAGATCTAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGAGCAGAAAGACCAATAAAGGGGCTCG--GAAGAACAGAAACAGCAGCAACCCAC
AGCACAAACAATAAAGAGAGGAOCTCAAAAGAGAACAGGACTGACCCAAACATCTC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCCCAGGCAGGTTTCTGGCTATAGTGCATCTCAGCTCTAACACAGCTGATAOCCCTT
TCCAGACCAAGATTCTGGCTTCACTACCCACTGCTATGACNACCTCTGCTCTNN

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGCTTTCAGTCAATTCOCCACGACAGATGATGGTGGAAATGTAAGTGCAGTTCCTATCA
NGCATTCACTCAITTCOCCACAGAAACACACACTGGGATTTCAAGCACATCAGTTTTCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTCAACAACCTCTCTCTGGCCA---TAACTGAACTGTTTGAG-AGGTAC--AOCAGAC
TTCAAGAACTCTTAATCTGACAGATGTGATTGAGAGACTAGCCACAGTACTCTAGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTGGGAATACACAGCTTTGGAAACAAGTTGTTGAGCAATCACAGGAGAGTACAC
TTTGAGAGCACAAATGCTTCTGAAACAACTTTGTCCAGCAATCACACAGAGTACAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGTGAAGAGGCTCA--GACAC--ACACCCCACTCTCAGCAGTGGGGGCCCCC
AACTAGGAAGCATATTAGACACTCAACCAACCACTTCTTGAGCAGCAGTCTACTCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGTGCOCACCTCCTTCCACCTCCTTTTACTAAGGGTGTGGTTACAGACAGCAAAGTCAC
AATGCCAGTTCCCATCTCCCTCCCTTTACTCAGAGAGCAGTTACTGACAACTGGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA
TOCCATTTCGGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCTCAAGGCACAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TACAGATCTGCAGCAACCTCAGCAGAGCTAGCCCAATCCTGAGATCATACTGGAAC
TOCMNNNNNNCA--AATGCCAAGTTCA--CNAATTTNGAACNNNNNACTCRNNN

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CACAGACTCTCCTCTAACTCTGTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATTA
NACTTCATCTACNTCTAATCTGTTACATTCTACTCCCATGCCAGCACTACACAGTTAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ACCACAGAATTCTAAATGGAAGCCCTCTCCTCGGCCAGACACAAATATCAGCTCAAGTC
ATCAGAGAATTCCAAATTAACCTCATCTCCTCGGCCAGAAATACCAATTTGGCACAAGCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATACTCGAAGAACCTTGAAGAGGGGCAAAAGGCCAGCACTAAGCATGTCCGCCACCTCAG
ATACTCAGACATTGCTGAAAAGGCCAAAAGGCCAGAGTAAGCATGTTGGCT-ACTACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

-CCTTCAGAGGCCAGCACTCATGCTCACAAGTGAATACACAGAAGCATGCAGAAAAGA
GCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAGAACACAAAGAGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTGTTTTTGATAAGAAACCTGGTCA--AACC--CAACTCCAAACATCTGCTTAAGTCT
GTGACTTTGATAAGAAACAGTTCAAGAGCAGACACTTCCAAACTCCTTCCCTTTGACT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTCTACCTAAGACTCTATTGAAAAAGCCAGATAATTGGAGGAAGGCTCCAGCTTTA
CTTTGTCTAGGTATATATTGAAAAGGCCAGGATAGTTGGAGGAAGAGCTGCAAGTTTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CAGTTCAGCTAATTCAGACGTTTTCTCTCTGTGGAGGCTGTGGAGAACCACTGCCCA
CTATTCCAGCTAACCTCAGATGCTTTCTTCCCTGTGAAGCTGTGGAAATCCCTGCCCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCATCCTACTGGACAGAGTTTCATCAGGANTTGAATATCCAGAGGACACAGAAAAGCC
CCATTCAATGGACAGAGTNNNNNTCAGGACTTGATTATCTAAGAGGAACAGAAATAGCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGTTCCAGCTGCTTCCCAATGGACCTTGTCCATCCAGAGGCTCAGTATTCAGGACCGTG
GGGTCCAGGTTCTCCCAATGGTACCTGTCCATCCAGAGGCTGGAATTCAGGACCGCG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GACAGTAACCTGTGCTCTGCATTTAATCCACTGGGCTAGAACATTTTCAGTCTCTTTGT
GACAGTACTTGTGTTCCGATCCAACTCTGTTGGCACAGACCACTTCATGTACCTTGT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGGTTTTTTTACCCGCAAGGATTTTGGACAGACATGTCAAGGAGATCAGTTTCACT
CTGTGGTTTCTATCTCTCCAGGATCTGGAGAGAGTACCAAGAGATCAGTTTCACT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TTGGAGTACTGTGGAACTAAAGTGCAGAGTGGAGGSTATGCGAGGCOCTACGGTTTCCT
COGGAAGCACTGTGGAACTGAGTGCAGAGCAGAGGTTAGGCOCAAGCOCTACAGTTAOCCT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGTACTTGCAAACCAACGGTGGTCTCAGAACGGCCAGGGGAGCAGAAAGGTCTGGG
GGATTCTTGCAAACCAACAGTTGTCTCAGATCATCCAGGGAGTAGGCAGGCTGTGG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA
TGACGGTTGACGGACATTGGTCTCCACAACTCTCAGTATTATGACCGTGGCTTTTACA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGATACAGTCATCA
AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCAGTGTGGTTAAGATACAGTCATTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGCTCCCCCTGTATTATAGAGCAAAAGAGGCAAGCCATCGTTGGGGTTTTAGGTGGAA
CAGCACACCTGTTATTCTAGAGCAAAAGAGGCCAGTCATTGTAGGCACCTTGGGGTGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTTTGAAACTGCCCTGCACTGCAAAAGGAAGTCCCGAGCCAGTGTTCAGTGGGTCTTT
GTTTAAAACTGCCCTGTACTGCAAAAGGAAGTCCCGAGCCAGGTTTACTGGGTCTCT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGATGGGACTGAACATAAAACCAATTGCAGTTGACTCAITCCAGATTTTCTGTATCCAA
CTGATGGCACTGAAGTGAAACCAATTACAGTTTACCAATTCAGATTTCTTATTTTCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGGGACTCTGTATATAGAAGCATCGCTCCTTCAGTGGGGGCATATAGAGTGCATTG
ATGGGACTTTGTATATAGAAGCACTAGCCTCTTCAGACAGGGGCATATAGAGTGCATTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCACAGCTCCTCAGGCTCAGAGAGAGGGTAGTGATTCTTACTGTGGAGAGGGAGAGA
CTACAGTTCACCTGGTTCGGAGGAGAGTAGTATGCTTACAAATGGAGAGGAGTAGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAATCCCCAGGATAGAACTGCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT
CCAGCCCCAGGATAGAACTGCTATCCAGAAAGGACTGAGGTGAATTTGGGTGAGAAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TACTACTGAAGTGTCTCAGTACTGGGGATCCAAAGCTAGAAATAATCTGGAGGCTGCCAT
TACTACTGAAGTGTCTCAGTACTGGGGATCCAAAGCTAGAAATAATCTGGAGGTTACCAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCAGGCTGTCTATGACCAAGTGGCCAGAAATGGGCAAGTCCCAATCCAGCTCTACCCAAATG
CCAGGCTGTGTGATGACCAAGTGGCCAGAAATGGGCAAGTCCCAATCCAGCTCTACCCAAATG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GATCCTTGGTGGTGGTTCAGTGACGGAAAAAGACGCTGGTGAATCTATGTGTGGCAA
GATCCTTGGTGGTGGTTCAGTGACGGAAAAAGACGCTGGTGAATCTATGTGTGTGGCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAACAAAAATGGGAGATGACCTAGTCTGATGATGTCCGCTGAGATTGACACCTGCCA
GAACAAAAATGGGAGATGATCTGATCTGATGATGTAGCTTAAAGACTGAAACCTGCCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAATTGAACAGAAGCAGTATTTTAAAGCAAGTGTCCATGGGAAAGATTTCAGTGTG
AAATTGAACCAAGCAGTATTTTAAAGCAAGTGTCCATGGGAAAGATTTCAGTGTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCCTGATGGGACAG
ATTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCCTGATGGGACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT
TGATCAACAATGCAATGCAAGCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCACAAATGGAACTTTGTATTTCACCAAGCTTGGGATGGCAGAGGAAGGAGATTATCT
TCAACAATGGAACTTTTATCTTCACCAAGTTGGGATGGCAGAGGAAGGAGATTATCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GCTCTGCCAGAACCTTAGGGAAAGATGAGATGAATGTCACCTAACAGTTCTAACAG
GCTATGCCAGAACCTTAGGGAAAGATGAATGAATGTCACCTAACAGTTCTAACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCATCCACGGATAAGGCAAGCTACAGACCAACATGAGGCTCAGGGCTGGAGAACAG
CTGCTCCCGGATAAGGCAAGTAAACAAACCAACAGAGAAATCAAAGCTGGAGACACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTCTTGACTGCGAGGTCAGTGGGAAACGAGGCCAATGTATTTGGTTGCTGCCTT
CTGTCTTGACTGTGAGGTCAGTGGGAAACGAGGCCAATGTATTTGGTTGCTGCCTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCAACAATGTCAATTCATTCTCAATGACAGGTCACATTTATGCCAATAGACTTTGT
CCAATGACATGATTTCTCTTCTCAATGATAGGTACACATTTATGCCAATAGGTCCTTGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCATCCATAAAGTGAAACCACTTGACTCTGGGACTATGTGTGCGTAGCTCAGATCCTA
CCATCAACAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCGAAATCCCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTGGGGATGACACTAAGACATACAACTGGACATTTCTCTTAACCTCCATTATCAATG
GTGGGGATGACACCAAAATGTACAACTGGATGTGTCTCTAACCTCCATTATCAATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GCCTGTATGCRAACAAGACTGTTATTAAAGCCACAGGCATTCGGCACTCCAAAAAATCT
GTCTGTATACRAACAAGACTGTTATTAAAGCCACAGCTGTGAGACATTCAAAAAATCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTGACTGCAGAGCAGATGGCATCCATCTTCAGGTCAGTGGATTATGCCAGGCAATA
TTGACTGCAGAGCTGAAGGACACCATCTCTGAAGTCATGTGGATCATGCCAGCAATA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTTCTCCAGCTCCATCTTTTGGAGCAGAGTCAGGTCATCCAAATGGAACCTTGG
TTTTCTCCAGCTCCATCTTTTGGAGCAGAGTCAGGTCATCCAAATGGAACCTTGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGATGAGGAACATCCGCTTTCTGACTCTGCGGACTTCACTGTGTGGTTCGGAGCGAGG
AAATTAGGAATGTGAGGCTTTCAAGTTCAGCGGACTTATCTGTGTGGGCAAAATGAG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT
GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCAGAAACCCATTCAACGAAAAAGTCATGGCCAGCTGGCAAGCCCGTAGCACTGAAC
TTAGAAATCCATTTAATGAAMATAAGTTGCCAGCTGGGAAAGTCCACAGCAATTGAATT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCTCTGTGATGGGAACCCCACTGAAATTAAGTGGATCTTACCTGACGGCACACAGT
GCTCTGTGATGGTAACCCACACCTGAAATAATCTGGATTTTACCAATGGCACACGAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TTGCTAACAGACCCACAAATCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT
TTCCAAATGGACCAAGATATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACAAAGCAACTGGGAACAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
CTAAACAACTCGGGAGGATGCAGGAAATATCGCTGTGCAGCTAGGAATAAGTTGGCT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACATCGAGAACTCATCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGAC
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGGATGGTGAAGAGCGTCAGTGGGGAACGTTATCACTGCATTGTGTCTCTGATGGGA
CAGGGACAGTAAGGAGCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTGATGGGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCCCCAAGCCAAATGTCRAAGTGGACTACACCGGGTGGCCATGTAAATGGACAGGCTCAAG
TCCCAAGCCAAATATCAAAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCTCAAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGGATGGAAATACATACTGCTGAAATGGCAAGCTGCTCATCAAGCAACACAGCTC
TTAATGGGAATACATATTGATGACAAATGGCACTTAGTCATTAAAGAGCAACAGCTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACGACCAAGGAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCAATATTAGCG
ATGACAGAGGAATATATCTGTAGGGCTCAAAATAGTGTGGTCTATACCTGATTACTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGTCAGTGTGTTGTGGCCATCCCTCCCGAATCATAACTAAGTACCCAGGACATGC
TTCCAGTAAATGATTGTAGCTACCTCCCGAATACAAATCGTCCACAGGAGTATTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCAGGAGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCAAGCCAA
TCACAGGACAGGGGAGCCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTCCCAAGCCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAGTCACTGGGAGACGCGAGACACTCCCTGCTCTCAAAAGCAACAGCAGAAACCCC
AAATCACATGGGAGTGGCTGACCACTCCCTTCTCTCAAGCGCAAGTAAGAGAGGAGAC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATAGAAGTGAGTGCCTTACCCCAAGGTACGCTGGTCAITTCAGAAATCCCAACCTGG
ATGGAAGTGAGAGCTTCACTTACAGGTACCTAGTCATTAGAAATCCCAACCTGG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATTCCGGAGTCTATAGTGCAGAGCTCAGAACCTACTTGGGACTGATTACGCAACACTT
ATTCTGGGATATACAAATGCACAGCAAGAACCACCTTGGTAGTGATTATGCAGCAACGT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACATCCGGTACTCTGACAGGAGGGGGAGACTAAAATTCAACAGAGTCCACATCCACA
ATATTCAAGTAATCTGACATGAA-----ATAATAAAGT-CAACAA-----CATCTGGGCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGGTTTATTTTGGGAAGAAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA
GAATTTATTTTGGGAAGAAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAATTTGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATACATTACAGTATTAAATTACAAATGGACATGCGA--TGA--GACTTGTAATGAAA
ATACATTACAGTATTAAATTACAAATGACATGCAAAATAAAGGACTTGTAAATAAAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCATTGTGAACGAA--ACCGAGTCTCTG--TGGATCTCAAAGCCTTCTTAACTTAA
GCATTATGAACGTATGATCTGATTTAATTTAATGGATCTCAAACAAACTTTTAACTTAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGCACTTTGATTTTGGCAACAAATAATAACAAACATTAAGAGAAAAAATGATCCACTAC
GGCACTTTTATTTTGGCAACAAATAATAACAA-----ACATTGAAACGGTTCACTAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GAAATAACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTCTCTCGCTAAC
AAAATAACAAATGGCTAATGTACCTGAATTTTTCAGTAAAAAA--TGAACCTT-CTAAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGTTGCCAGCTGCTCTGTCTCTTCTACCAATGTCAAAACATCGCACACAGGGTGA
A-----CCAGTTGCCTAGTGTCCACTCTCTATCAATGTACAAAGCATGGCACTCAG-----A

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATGGAGTCAACGGGAAGATTAAAGTTTGGGTCGTGTAAATCTCAATGTACAAATATTC
ACAGAGACAAATGGAAAAATTAATCTGC-----AATCT--ATGTATAAATATTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGTCTCTGGTTTATAAACATTTT-GATAAAACCGAAAAAATAAATAAATAAATAAATAA
TGT-----GGTTTATAAATTTTGTCTAAACCTACAGAAATAAG-----

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAA

(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVGRGVSGLLISLTAVCLVVTGSRACPRRCACVPTVHCTFRYLTSIPDGIPANVE
human_5+3_corrected	MKVGRGITCLLVSEFAVICLVATPGGEACPRRCACVPTVHCTFRYLTSIPDSIPNVE
mouse_5_corrected	MQKRGREVSCLLISLTATCLVVTGSRVCPRCACVPTVHCTFRDLTSIPD-GPANVE
rat	RINLGYNLSLTALTENDFGLSKLELIMLHSHNGIHRVSDKTFSGLSQVLKMSYKVKQII
human_5+3_corrected	RINLGYNLSLVIMETDFSGLTLELIMLHSHNGIHTIPDKTFSDLQALQVLKMSYKVKRL
mouse_5_corrected	RVNLGYNLSLTALTENDFSGLSKLELIMLHSHNGIHRVSDKTFSGLSQVLKMSYKVKQII
rat	RKDTFTGLGSLVRLHLDENNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
human_5+3_corrected	QKDTFYGLRSLTRLHLDENNIEFINPEFVYGLNTRLRLVHLEGNQLTKLHPDTFVSLSYLQ
mouse_5_corrected	EKDTLYGLRSLTRLHLDENNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQ
rat	IFKTSFIKYLELSDNFLTSLPKENVSYMPNLESYLYHGNPWTCDCHLKNLSEWMOGNPDI
human_5+3_corrected	IFKISFTKFLYLSINFLTSLPQENVSYMPDLDSLYLHGNPWTCDCHLKNLSDWIQ--PDV
mouse_5_corrected	IFKTSFIKLYLYDNF-TSLPKENVSSMPNLESYLYHGNPWTCDCHLKNLSEWMOGNP--
rat	IKCKKDRSSSSPQCCPLCMNPRI SKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTOEDNG
human_5+3_corrected	IKCKKDRSPSSAQCCPLCMNPRTSKGKPLAMVSAARFQCAKPTIDSLKSKSLTILEDS
mouse_5_corrected	-----
rat	SASTSPQDFIEPFGSLSLNMTXISGNKADMVCSIQKPSRTSPATTEENDYIMNASPST
human_5+3_corrected	SAFISPGGFMAPFGSLTLMNTDQSGNARMVCSIQKPSRTSPATTEENDYIVLNTSPST
mouse_5_corrected	-----
rat	NLVCSVDYNIHQPVWQILALYSDSPLILERSKPLTETPSLSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIQYGHQPVWQILALYSDSPLILERSHLLSETPQLYYKYQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADFFWFQCKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTIMLMMN
human_5+3_corrected	DLRADPSWLMQDQISLQLNRTATTLSTLQIQYSSDAQITLPRAMRPVKHWTIMLSDNN
mouse_5_corrected	-----
rat	PKLERTVLVGGTIALSCPGKGDPSPHLENLLADGSKVRAPYVSEDGRILIDKNGKLELQ
human_5+3_corrected	TKLEHTVLVGGTIVGLNCPGQDPTPHVDWLLADGSKVRAPYVSEDGRILIDKNGKLELQ
mouse_5_corrected	-----
rat	ADSFDAGLYHCISTNDADADVLTIRITVVEPYGESTHDSGVQHTVVVIGETLDLPCLSTGV
human_5+3_corrected	ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGI
mouse_5_corrected	-----
rat	PDASISWILPGNTVFSQPSRDRQILNNGTLRILOVTPKQGHYQCVAAANPSGADFESFKV
human_5+3_corrected	PDASISWVTPGNVLYQSSRDKKVLNNGTLRILOVTPKQGGYTRCVAANPSGVDFLIPOV
mouse_5_corrected	-----
rat	SVQKKGQRMVEHREAGGSGLGEFNSSVSLKQPAKLSASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVKMGQRFLEHDGETEGSGLDESNIPIAHLEKPPGAQLRTSALMEAEVGGHTSTGKRRN
mouse_5_corrected	-----
rat	HRDLIHRRGDSTLRRFREHRRQLPLSARRIDPQWAALEKAKKNSVPPKQENTTVKPV
human_5+3_corrected	YRELTLORRGUSTHRRFRENRKHFPPSARRIDPQWAALEKAKKNSVPPKQENTTVSPF
mouse_5_corrected	-----

rat PLAVPLVELTDEEKDASGMIPDEEFMLTKASGVPRSPADSGPVNHGEMTSIASGT
human_5+3_corrected PVVTQLFNIPGEEDOSSGMIALHEEFMVPAKALNLPARTVTADSRITSDSPMTNINYGT
mouse_5_corrected -----

rat EVS-TVNPQTLOSEHLPDFKLFVINGTAVTKSMNPSIASKIEDTINQPIIIFF---SV
human_5+3_corrected EPSFVVNSQILPPEEPTDFKLSTAIKTAMSKNINPTMSSQIQGITNQHSSVTFPLLGA
mouse_5_corrected -----

rat AEIRDSAQAGRAS---SQSAHPVTGGMATYGHINTYSSFTSKASTVLQPINPTESYGPQI
human_5+3_corrected TEFQDSQMGGRGHEHFQSRPPITVTRMIKDVNVKMLSSTTNKL---LLESVNTTNSH--QT
mouse_5_corrected -----

rat PITGVSRPSSSDISSHTADPFSSEHPSGSHTTASSLFHIPNNNTGNFPLSRHLGREIT
human_5+3_corrected SVREVSEPRNHFYSHTTQILSTSTFPSPDPTAHSQFPPIRNS-TVNIPLFRFRGRORR
mouse_5_corrected -----

rat IWSRGVKNPHRTFVLRHRHRTVPAIKGPANKNVSVQVATEYPMCHTCPSAEGTLVA
human_5+3_corrected IGGRGRIISFYRTFVLRHRHYSIFASTTRGSSEKSTTAFSATVLMVTCLSLPRERLTA
mouse_5_corrected -----

rat TAALSVPSSESSALPKTNVGVIAZESTTVVKPILLFKDKQNVDEIIITTTKYSGGES
human_5+3_corrected TAALSFPSAAPITFPKADIARVPSEESTTLVQNPILLLENKP--SVEKTPPTIKYFRTEI
mouse_5_corrected -----

rat NHVIPTASMTSAPTSVSLGKSPVDNSGHLMPGTIOTGKDSVETTELPSPLSTP--SIP
human_5+3_corrected SQVPTTGAVMTYAPTSTIMEKTHKVNASYPRVSSVTEAKRDSVITSSLSGAITKPMTHI
mouse_5_corrected -----

rat TSTKFSKRRTPLHQIFVNNQKKEGMLKNPYQFGLQKNFAAKLFXIAPLLPTGQSSPSDST
human_5+3_corrected AITRFSRRTIPWQONFVNNHNPGRRLRNQHKVSLQKSTAVMLPKTSPALPQROSSPFHFT
mouse_5_corrected -----

rat TLLTSPFPALSTMAATONKGTETVVGARSLGAKQ-PYTNSSPVLPSSTIKRSNTINF
human_5+3_corrected TISTSVMQIPSENTLTAHHTTTKTHNPG-SLPT-KKELPFPPLNPLPSIISKDSSTKSI
mouse_5_corrected -----

rat ISTETPT-VTSPTATASVIMSETQRTSKEAKDQIKG-F-RNNENNANTTPROVSGYSAY
human_5+3_corrected ISTQTALPATTPFPASVITYETQTERSAQTIQREQEPQKERTDPNISPDQSSGFTTP
mouse_5_corrected -----

rat SALITADTFLAFSHSPRODDGCVNSAVAYESTTS--LLAITELFEKITTQTLGNTTALETT
human_5+3_corrected TAMTP--PALAFTHSPPENTTIGISSTISFHSRTINLTDVIEBLAQASTQTLKSTIASETT
mouse_5_corrected -----

rat LLSKSQESTTVKRAS-DTP-PFLLSAGAPPVETPSPPPFETKGVVTDGKVTSAFQMTSNRV
human_5+3_corrected LSSKSHQSTTTTKASLDTPIPPFLSSSATIMPVPISPFTQRAVTDTRGDSDHFRIMNTV
mouse_5_corrected -----

rat VTIYESSRHNTDLOQPASAEASPNPEIITGTTDSPNLFPSTSVPALRVDPKONSKWKESP
human_5+3_corrected VKLHESSRN--LQMPSSQLEP-----LTSSTSNLLHSTPMPALITTVKSONSKIATPSP
mouse_5_corrected -----

rat WPEHKYQLESYSETIEKGRPAVMSMSPHLSLEASTHASHWNTQKHAESVFDKPKGQNP
human_5+3_corrected WAEYQFWHKPYSIDAEKGRPEVSMIATTGLSEATTLVSDWDGQNTKASDFDKKPVQEA
mouse_5_corrected -----

rat -TSKHLPPYVSLPKTLKKPRIIGKKAASFTVPANSDFVLPCEAVGOPLPIIHWTRVSSGX
human_5+3_corrected TSKLLPFDLSRYIFEXPRIVGGKAASFTIPANSDAFLPCEAVGNELPTIHWTRVS-GL
mouse_5_corrected -----

rat KISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFFVLSVVFYFARILD
human_5+3_corrected DLSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCSASKLFGTDHLHVTLVSVVYPPRILE
mouse_5_corrected -----

rat RHVKBITVHFGSTVELKCRVEGMPRTTVSWILANQTVVSETAKGSRKVVWTPDGTLLIYN
human_5+3_corrected RRTKEBITVHSGSTVELKCRAGCRPSPTVTWILANQTVVSESQGSRAQVVTVDGTLVLHN
mouse_5_corrected -----

rat LSLYDRGFYKCVASNPGGQDSLLVKIQVITAPPVITTEQKROAIVGVLGSSKLKPCTANGT
human_5+3_corrected LSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVITLQRRQVIVGTWGESKLKPCTANGT
mouse_5_corrected -----

rat PQPSVHWVLYDGTTELKPLQLTHSRFFLYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV
human_5+3_corrected PQPSVYVWVLSDGTVEVKELQFTNSKLFLFSNGTLYIRNLASSIRGTYECIATSSSGSERRV
mouse_5_corrected -----

rat VILTVEEGETIPRIETASQKNTVNLGEKLLNCSATGDPKPRIIWRLPKAVIDQWHRM
human_5+3_corrected VMLTMEERVTSPIEASQKRTVNFPGDKLLNCSATGEPKQIMWRLPKAVVDQ-----
mouse_5_corrected -----

rat GSRHVYPNGSLVVGSVTEKDAGDYLCAVARNKMGDDLILMHVRLRLTPAKIEQKQYFKQ
human_5+3_corrected GSNHVYPNGSLFIGSVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDKQYFRKQ
mouse_5_corrected -----

rat VLEGKDFQVDCRAGSGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNCTLYFNWV
human_5+3_corrected VLEGKDFQVDCRAGSGSPVPEISWSLPDGTNINNAMQADDSGHRTRRYTLFNNCTLYFNKV
mouse_5_corrected -----

rat GVAEEGDYTCYAQNTLGKDEMKVHLTVITAIPIRQSYKTMRLRAGETAVIDLCEVTGEP
human_5+3_corrected GVAEEGDYTCYAQNTLGKDEMKVHLTVITAIPIRQSNKTKRIKAGETAVIDLCEVTGDP
mouse_5_corrected -----

rat KPNVFWLLPSNNVISFSNDPFTFANRRLSIRHKVKPLDSGDYVCVAQNPSGDDTKYKLD
human_5+3_corrected KPKIFWLLPSNDMISFSIDRYTTFHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYKLD
mouse_5_corrected -----

rat IVSKPPLINGLYANKTVIKATAIRHSKXYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSR
human_5+3_corrected VVSKPPLINGLYTNRTVIKATAVIRHSKXHFDCRAEGTPSPPEVMWIMPONIFLTAPYFGSR
mouse_5_corrected -----

rat VTFVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPINEKVIA
human_5+3_corrected ITVHKNGTLEIRNVRVLSDSADFTCVARNEGGESVLVVQLEVLEMLRRPTFRNPINEKIVA
mouse_5_corrected -----

rat QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPNSPYIMAGNGSLILYKATRNKSGKY
human_5+3_corrected QLGKSTALNCSVDGNPPPEITWILPNGTRFSNGPQSYQYLIASNGSPTIISKTTREDAGKY
mouse_5_corrected -----

rat RCAARNKVGYIEKLILLEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDDGIPKPNVWHTTP
human_5+3_corrected RCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGBSLSLHCVSDDGIPKPNKWHTMP
mouse_5_corrected -----

rat GGHVIDRPQVDGKYILHENGTLVIKATTAHDOGNYICRAQNSVGQAVIEVSVMVVAIRPK
human_5+3_corrected SGYVVDRPQINGKYILHENGTLVIKEATAYDRGNYICRAQNSVGHTLITVPVMIVAYPPR
mouse_5_corrected -----

rat IINYLERNMLRRTGEANOLHCVALGIPKPKVTWETPRHSLLSKATARKPRSEMLHPQGT
human_5+3_corrected ITNRPPRSIVTETGAAPQLHCVALGVFKPEITWEMPDHSLLSTASKERTHGSEQLHLQGT
mouse_5_corrected -----

rat LVIQNLQTSDSGVYKCRQNLGTDYATTYIQV
human_5+3_corrected LVIQNPQTSDSGIYKCTAKNPLGSDYAATTYIQV
mouse_5_corrected -----

(rat: SEQ ID NO:10)

(human_5+3_corrected: SEQ ID NO:11)

(mouse_5_corrected: SEQ ID NO:12)

Figure 16

rat	MOVGRGVBGLLISLTAVCLVVTFGSRACPRRCACVYVTEVECTFRYLTSPDGIPANVE
human_5+3_corrected	MKVGRGITCLLVSTAVICLVATPGGRACPRRCACVYVTEVECTFRYLTSPDGIPANVE *:*** :: **:*.:.:***.***:*****;*****;*****.***.***
rat	RINLGYNSLTRLTENDFDGLSKLELLMLHNSGIERVSDKTFSGLOSLQVLKMSYNKVQII
human_5+3_corrected	RINLGYNSLRLMETDFSGLTLELLMLHNSGIEHTIPDKTFSDIQALQVLKMSYNKVRKL *****.***.***.***:*****;*****.***.***:*****;*****.***.***
rat	RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLEPDTFVSLSYIQ
human_5+3_corrected	QKDTFYGLRLSLRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLEPDTFVSLSYIQ :*****.***.***:*****;*****.***.***:*****;*****.***.***
rat	IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLGHNPTWCDCHLKWLSWNOGNPDI
human_5+3_corrected	IFKISFIKYLFLSDNFLTSLPKEMVSYMPNLESYLGHNPTWCDCHLKWLSWNOGNPDI ***.***.***:*****;*****.***.***:*****;*****.***.***
rat	INCKKDRSSSSPQQCPICMNPFRISGRPFAMVPSGAFLCTKPTIDPSLKSLSVTQEDNG
human_5+3_corrected	INCKKDRSPSSAQCPICMNPFRISGRPFAMVPSGAFLCTKPTIDPSLKSLSVTQEDNG *****.***.***:*****;*****.***.***:*****;*****.***.***
rat	SASTSPQDFIEPFGSLINMTXSGNKADMVCSIQKPSRISPTAFTEENDYIMLNASTST
human_5+3_corrected	SAFISPPQGFMAPFGSLINMTDQSGNEANMVCSIQKPSRISPTAFTEENDYIVLNTSFT **.***.***:*****;*****.***.***:*****;*****.***.***
rat	NLVCSDVYNHIIQPVWQALLYSOSPLILERKQLTETPSLSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCHIDYGHIIQPVWQALLYSOSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEA ***.***.***:*****;*****.***.***:*****;*****.***.***
rat	DVRADPFWFQOEKIVLQNLRTATLSTLIQIFSTDAQIALPRAEMRAERLKWMTILMMNN
human_5+3_corrected	DLRADPFWLMQDQISLQNLRTATLSTLIQIFSTDAQIALPRAEMRAERLKWMTILMMNN *:***.***:*****;*****.***.***:*****;*****.***.***
rat	PKLEHTVLVGGTIALSCPGKGDPSPHLEWILADGSKVRAPYVSEDGRILIDKNGKLELQ
human_5+3_corrected	TKLEHTVLVGGTVGLNCPGQGDPTPHVWLLADGSKVRAPYVSEDGRILIDKNGKLELQ :***.***:*****;*****.***.***:*****;*****.***.***
rat	ADSFAGLYHCISTNDADADVLITRITVVEPYESTHDSGVQHTVVTGETLLDLFCLSTGV
human_5+3_corrected	ADSPDTGVYHCISLNDADADVLITRITVVEPYESTHDSGVQHTVVTGETLLDLFCLSTGV *****.***.***:*****;*****.***.***:*****;*****.***.***
rat	PDASISWILPGNTVFSQPSRLRQILNNGTLRLILQVTEKDGHYQCVANPAGDPSSFKV
human_5+3_corrected	PDASISWVIFGNVLYQSSRLRQILNNGTLRLILQVTEKDGHYQCVANPAGDPSSFKV *****.***.***:*****;*****.***.***:*****;*****.***.***
rat	SVQKKGQRMVHEDREAGSGLGEPMSSVSLKQFASLKLASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVKMKGRPLEHDETEGSGLGEPMSSVSLKQFASLKLASALTGSEAGQVSGVHRKNK *:***.***:*****;*****.***.***:*****;*****.***.***
rat	HRDLIHRRGDSTLRRFREHRRQLPSARRIDPQWAALEKAKKNVSVPKQENTIVKPV
human_5+3_corrected	YRELTLQRRGDSTLRRFREHRRQLPSARRIDPQWAALEKAKKNVSVPKQENTIVKPV :***.***:*****;*****.***.***:*****;*****.***.***
rat	PLAVPLVELTDEEKDASGMIPPESEFVLTAKASGVPGRSPTADSGPVNHHGEMTSIASGT
human_5+3_corrected	PVVTQPLNIPGEEDDSGMLALHEEFNVFATKALNLPARTVTDGSRISDSFMTNINYGT *:***.***:*****;*****.***.***:*****;*****.***.***
rat	EVS-TVNPQTLQSEHLPDFKLSVINGTAVTKSMNPSIASKIEDTTNQNPIIIFP---SV
human_5+3_corrected	EFSPVNSQILPPEPTDFKLSTAKKTAMSKNINPTMSSQIQGTTNQHSSVTFPLLGA *.***.***:*****;*****.***.***:*****;*****.***.***
rat	ABIRDSAQAGRAS---SQSAHFVTGGNMATYGTNTYSSFTSKASTVLQPINPTESYGRQI
human_5+3_corrected	TEFQSDQNGRGRHEFQSRPITVTRIMKDVNVMKLSSTTNKL---LLESVNTTNSR---GT *:***.***:*****;*****.***.***:*****;*****.***.***

[illegible]

rat	SRIHVYPNGSLVVGSVTERDAGDYLCVARNKNGDOLVIMHVRLRLTPAKIBQKQYFKQV
human_5+3_corrected	SWIHVYPNGSLFIGSVTERKSGVYLCVARNKNGDOLIMHVSLRLTPAKIDBKQYFRQV *****;*****;* *****;**** **;****;****;***
rat	LHGKDFQVDCASGSPVPEVSWSLPDGTVLNNVAQADDSCGYRTKRYTLFNNGTLYFNNVG
human_5+3_corrected	LHGKDFQVDCASGSPVPEISWSLPGDTMINNAMQADDSCGRTTLYTLFNNGTLYFNNVG *****;*****;***. *****;****;*****;****;***
rat	MAEGDYICSAQNTLGKDEMKVHLTVLTAIPRIQSYKTIMRLRAGSTAVLDCEVTGEPK
human_5+3_corrected	VAEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTNKRIKAGDTAVLDCEVTGDEK ***** * *****;*** ***** **; **;****;*****;***
rat	PNVFWLLPSMNVISFSNDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSSGDDTKTYKLDI
human_5+3_corrected	PKI FWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSSGDDTKMYKLDV **;*****;**** **;***** **;*** *****;***** *****
rat	VSKPPLINGLYANKTVIKATAIRHSKQYFDCRADGTFSSQVTWIMPONI FLPAFYGSRV
human_5+3_corrected	VSKPPLINGLYNTVTIKATAVRHSKQYFDCRAEGTSPSEVMWIMPONI FLTAPYGSRI *****;****;*****;*****;****;*** *****;*****;****;***
rat	TVHNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVAQ
human_5+3_corrected	TVHNGTLEIRNIRLSDSADFTCVARNEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVAQ *** ****;****;***** **;*****;*****;*****;****;***
rat	AGKPVALNCSVDGNPPPEITWLLPDGTQFANRPHNSFYLMAGNSLILYKATRNKSGKYR
human_5+3_corrected	LKSTALNCSVDGNPPPEITWLLPMTGTRFSGNQSQYLIASNGSPITSKYTREDAGKYR **;***** ***** ****;***;*** **; ****;****;***;****;****
rat	CAARNKVGYTEKLILLEIGQKPVILTYEPMVKSVSSEPLSLHCVSDGIPKPNVKWITPG
human_5+3_corrected	CAARNKVGYTEKLIVILEIGQKPVILTYEPMVKSVSSEPLSLHCVSDGIPKPNIKWITPS *****;***** ***** **;*** *****;**** *
rat	GEVIDRQVQDGKYLHENGTLVLIKATTAHDQNYICRAQNSVGQAVLSVSVNVVAYPPRI
human_5+3_corrected	GYVDRPQINGKYLLHDNGTLVKEATAYDRGNYICRAQNSVGHITLITVPVMIVAYPPRI *;****;*****;***** **;****;*****;*****;****;****
rat	INYLFRMLARTGEAMQLHCVLGIPIPKVYMETPRHSLLSKATARKPRSEMLHPQGTI
human_5+3_corrected	TNRPPASIVTRTGAFLHCVLGVKPEITWEMPDHSLSTASKERTHGEQLHLQGTI * **;*** *****;****;*** *****;*** **;*** *****
rat	VIONLOTSDSGVYACRAQNLGTDYATYIQVL
human_5+3_corrected	VIONPQTSDSGIYACTARNPLGSDYATYIQVI **** *****;*** **;*****;*****;

(rat: SEQ ID NO:13)

(human_5+3_corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLSQLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGNPWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPPNVE
RINLGYNLVRMLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
QIFKISFIKFLYLSDNFLTSLPQEMSYMPDLSLLYLHGNPWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAMRPVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTG VYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGIP
DASISWVIPGNVLYQSSRDKKVLNNGTLRLQVTPKDQGYRCAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRH
NYRELTQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALLEKAKKNAMPDKRENTTV
SPPPVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSR TISDSPMTNIN
YGTEFSPVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG
ATEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFPSPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGRGRIS
PYRTPVLRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPVSVSTNEAKRDSVITSSLSGAIKPPMTIIAITFRSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNLTLT
AHHTTTKTHNPGSLPTKKELFPPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETTLSSKSHQSTTTRKASLDTPIPPFLSSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLQMPSSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGGKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDLSRYIFEKPRIVGGKAASFIPANSDAF
LPCEAVGNPLPTIHWTRVSGDLDSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCSASNL
FGTDHLHVTLSVVSYPRIERRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNLISIYDRGFYKCVASNPGGQDSSLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPTAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSSTGSERRVVMLTMEERVTSPIEAASQKRTEVNFGDKLLLNCSATGEP
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTN
 KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVSKPPLINGLYTNRTVIKATAVRHSHKHFDCRAEGTP
 SPEVMWIMPDNIFLTAPYYGSRTVHKNGTLEIRNVRLSADFICVARNEGGEVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEIHWLPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVGYIEKL VILEIGQKPVILTYAPGTVKGISGE
 SLSLHCVS DGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGNYICKA
 QNSVGHTLITVPVMIVA YPPRITNRPPRSIVTRTGAA FQLHCVALGVPKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

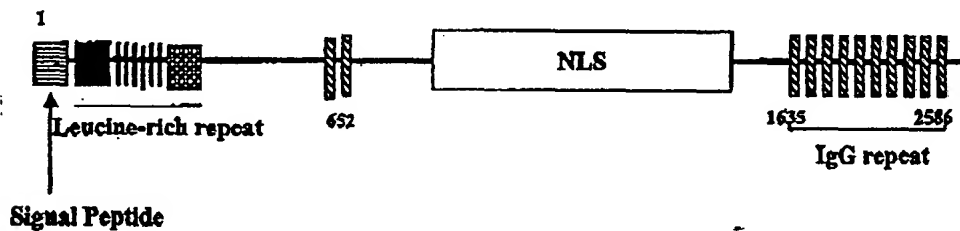


Figure 20

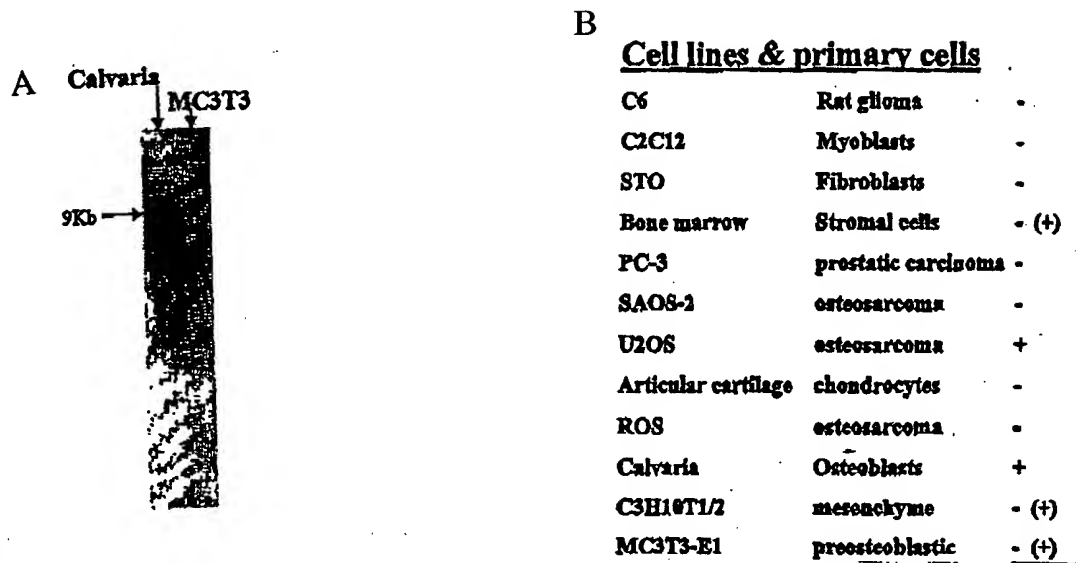


Figure 21

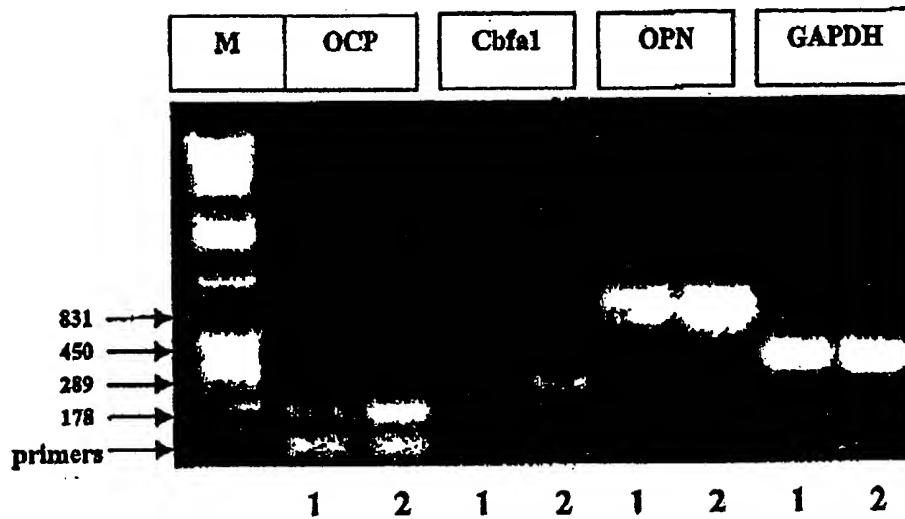


Figure 22

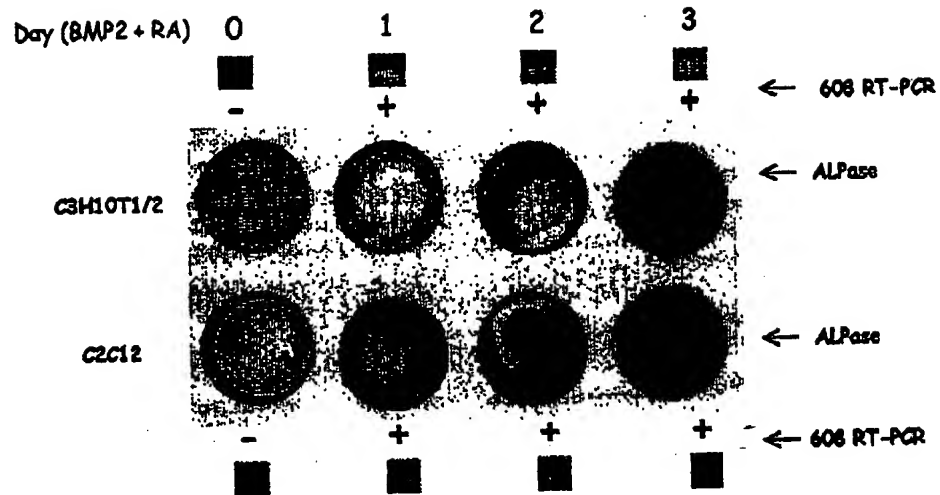


Figure 23

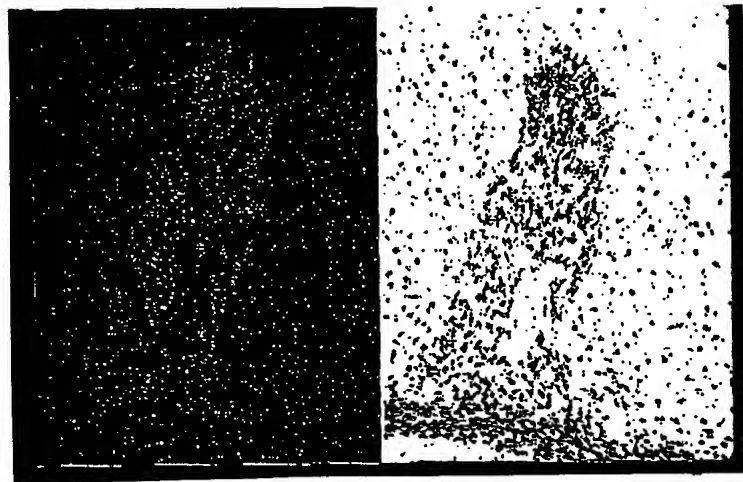


Figure 24

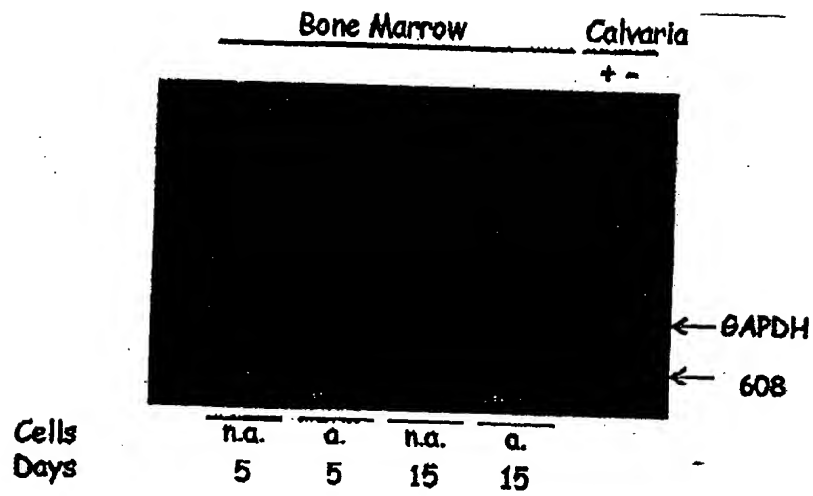


Figure 25

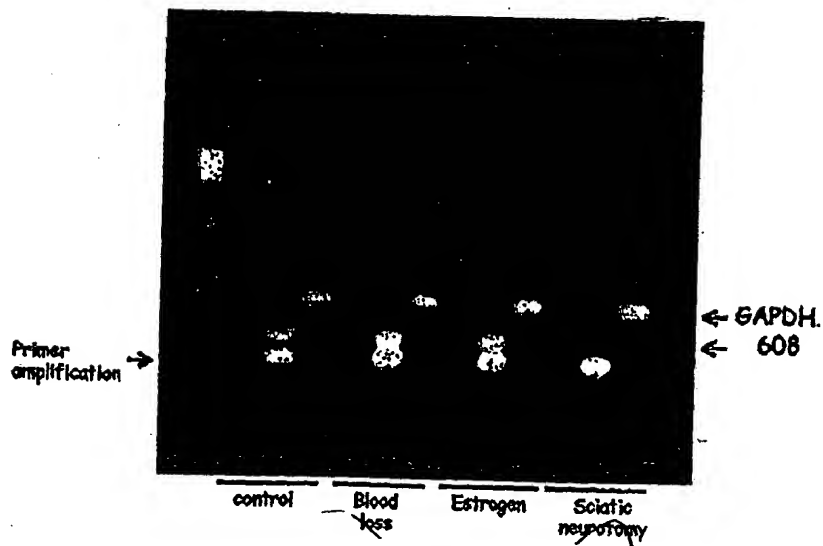


Figure 26



Figure 26 is a photograph of a vertical, textured object, possibly a piece of wood or a biological specimen, showing a central lighter area and dark, irregular edges.

Figure 27

A



B



Figure 28

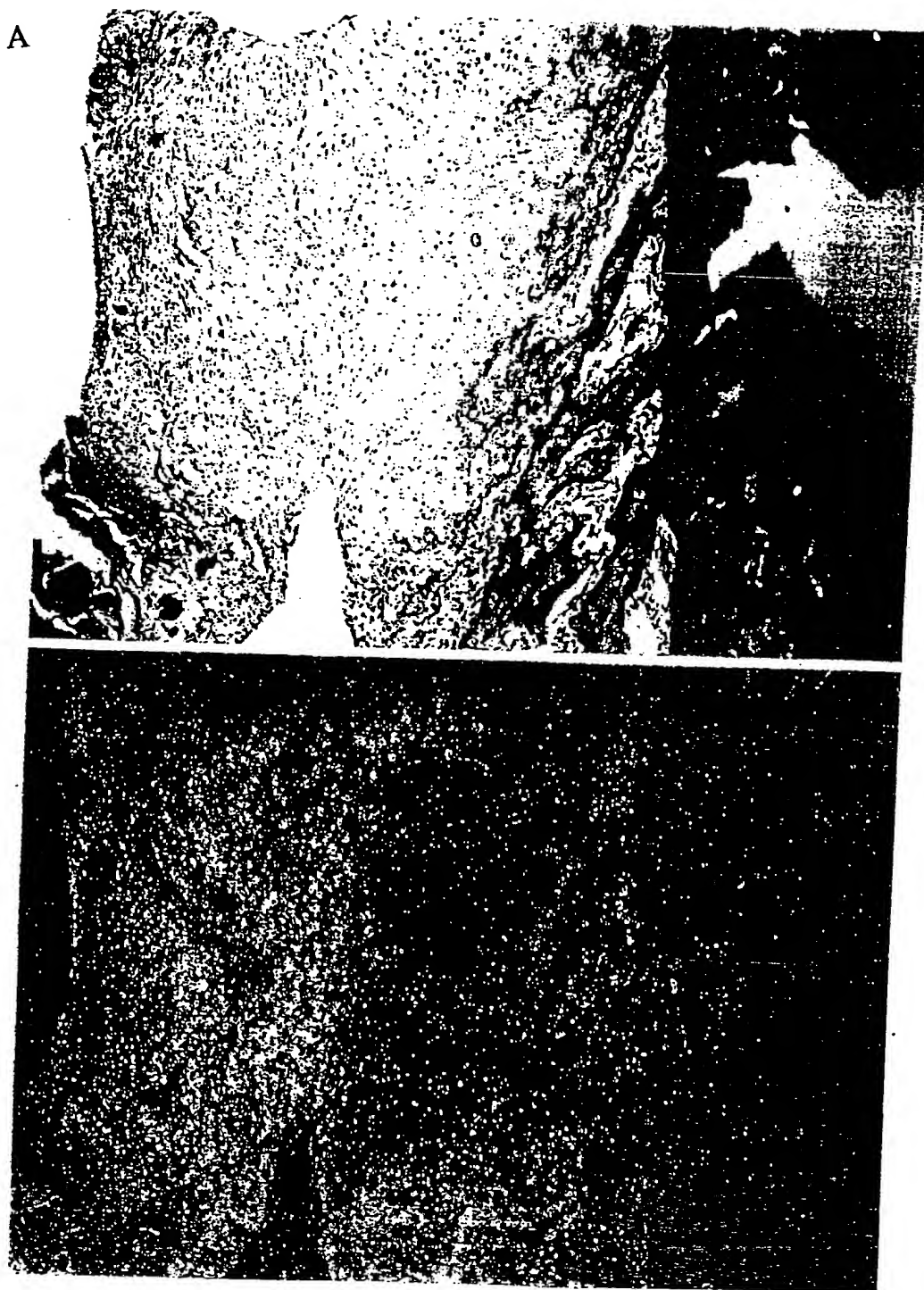


Figure 29

A



B

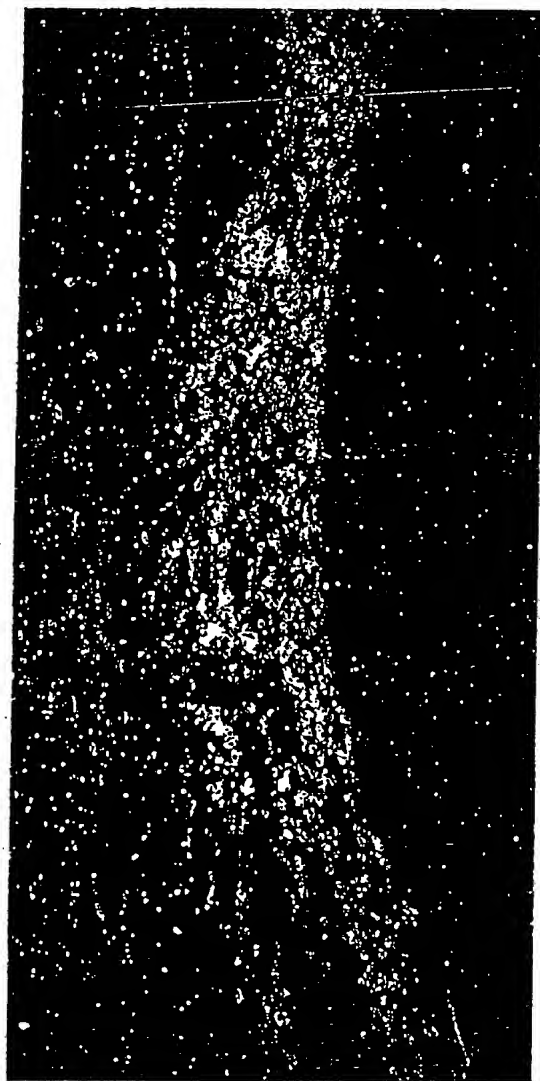


Figure 30

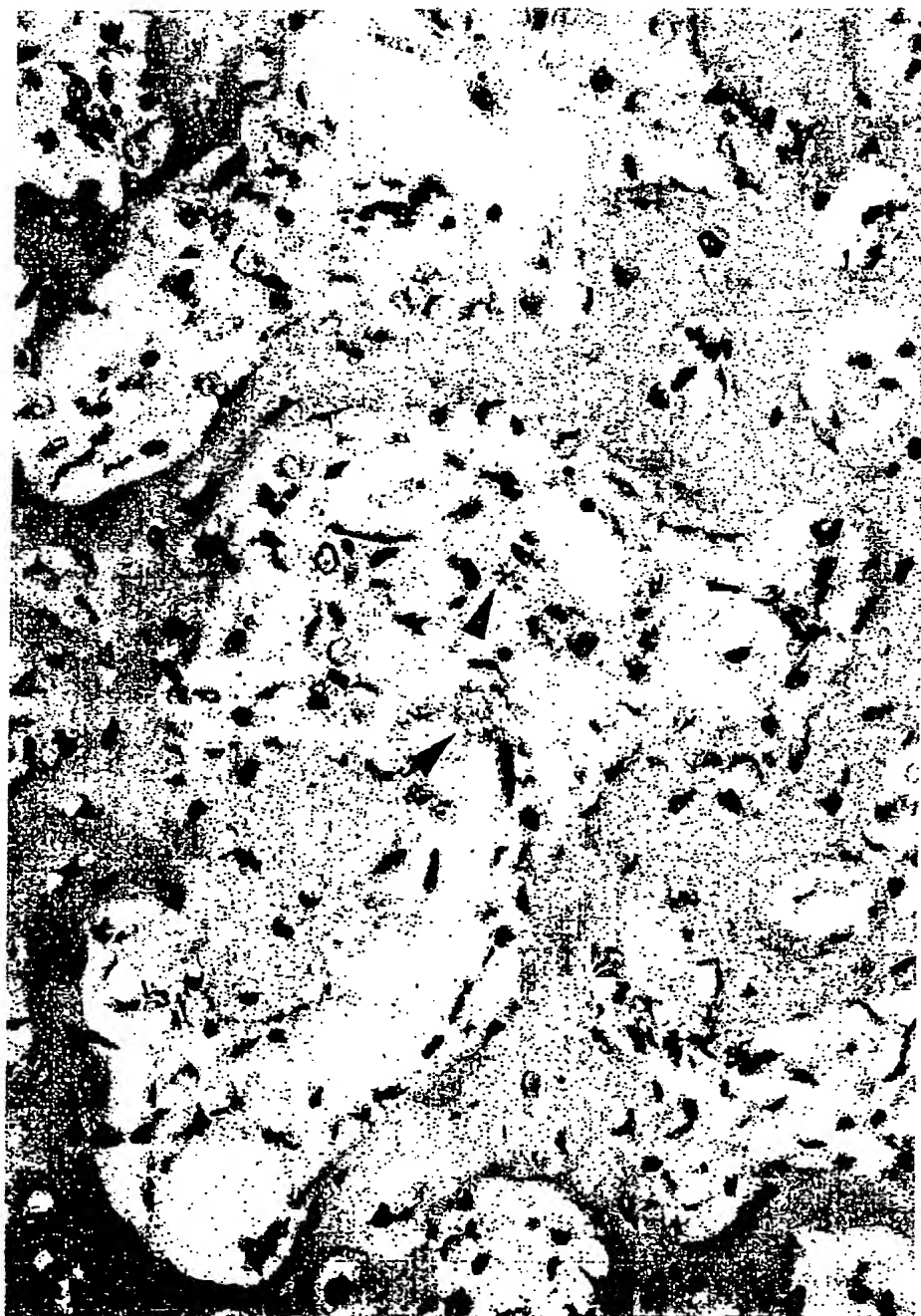


Figure 32

A



B

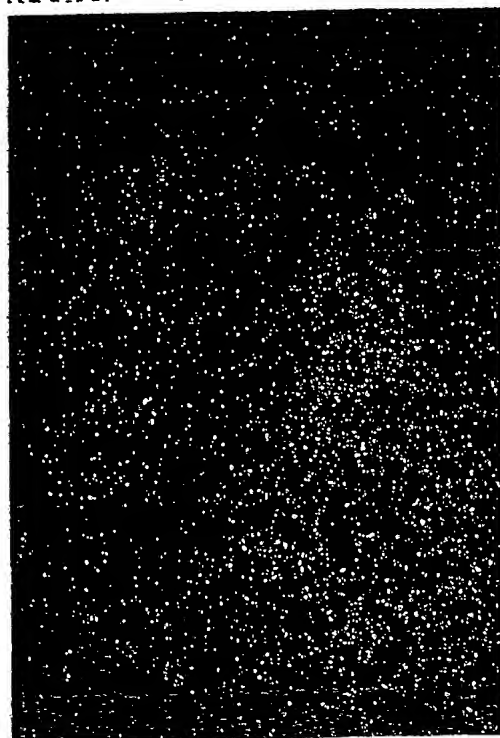


Figure 33

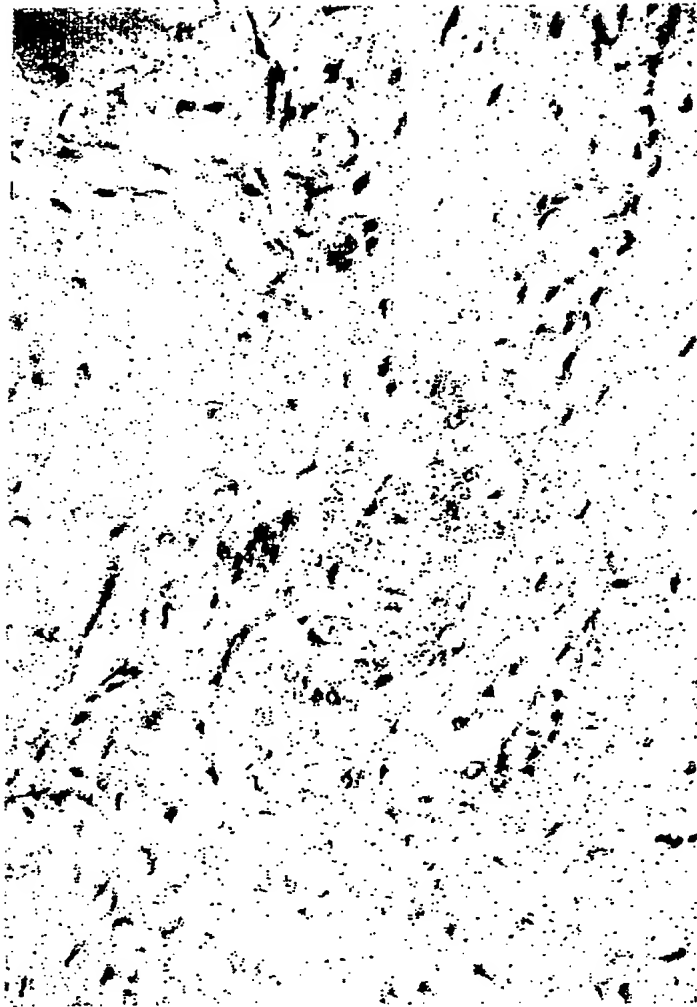


Figure 33

Figure 34

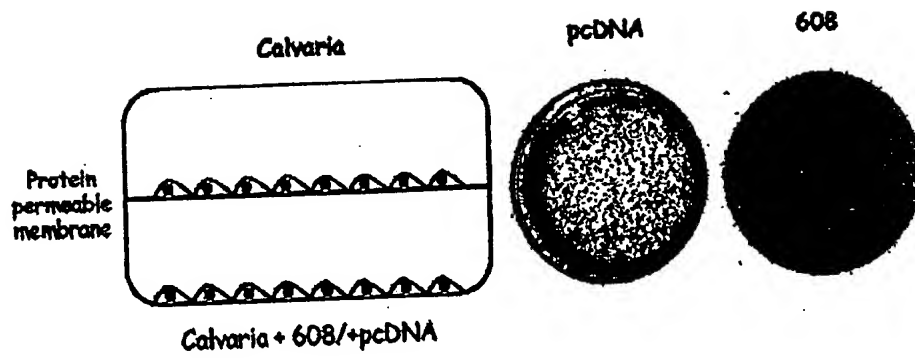


Figure 35

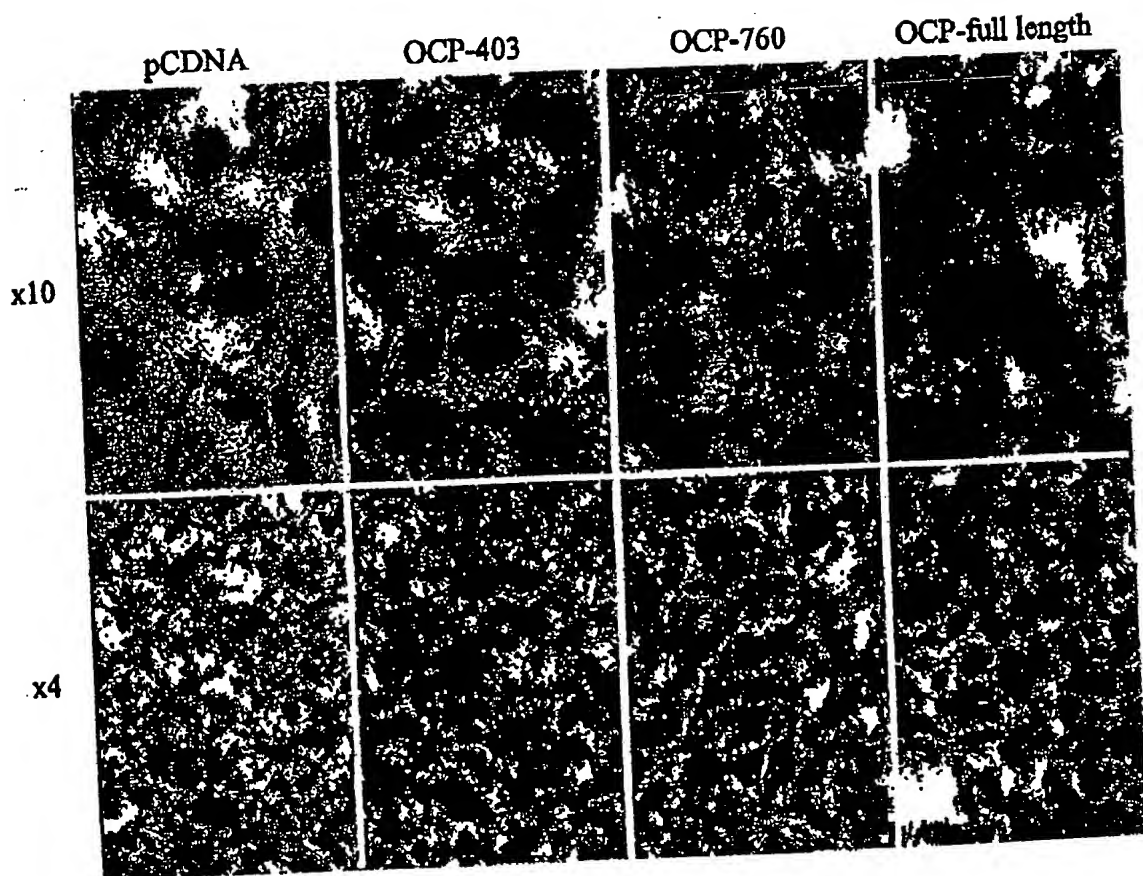


Figure 36

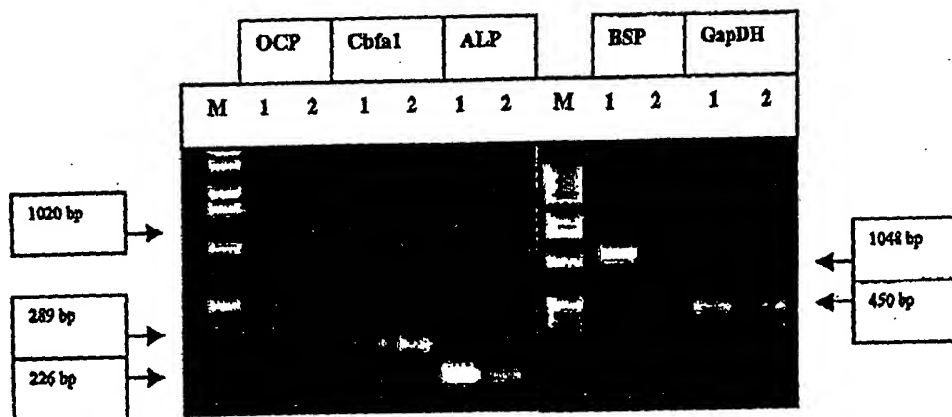


Figure 37

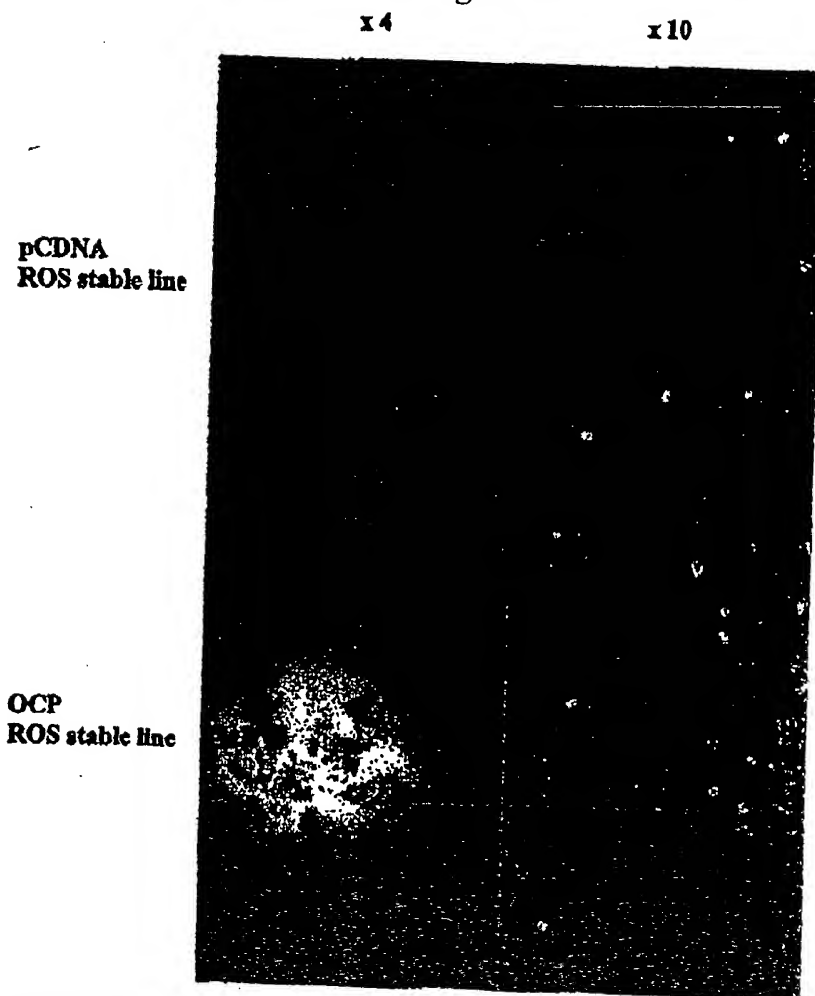


Figure 38

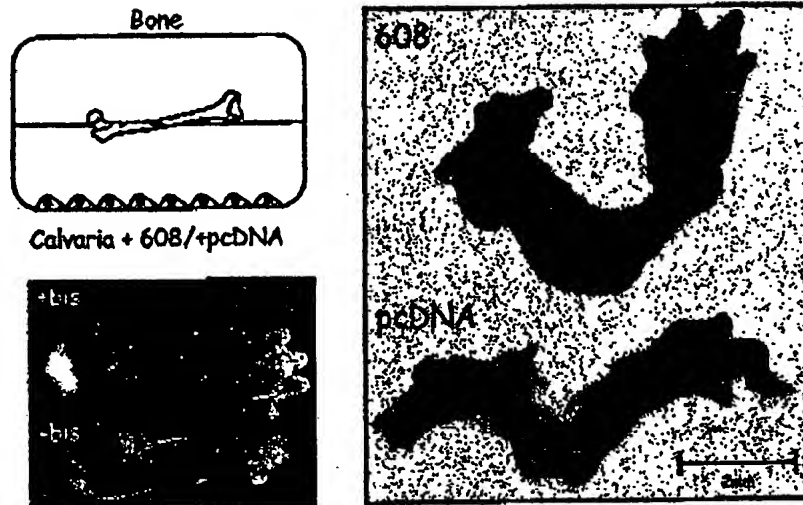


Figure 39

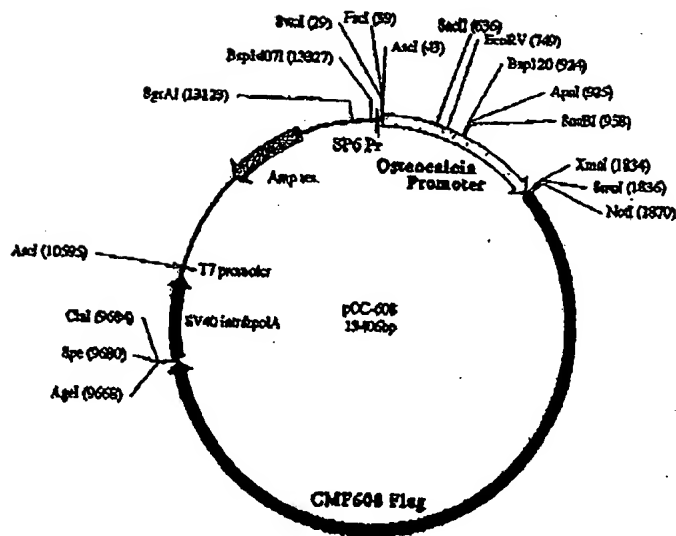


Figure 40

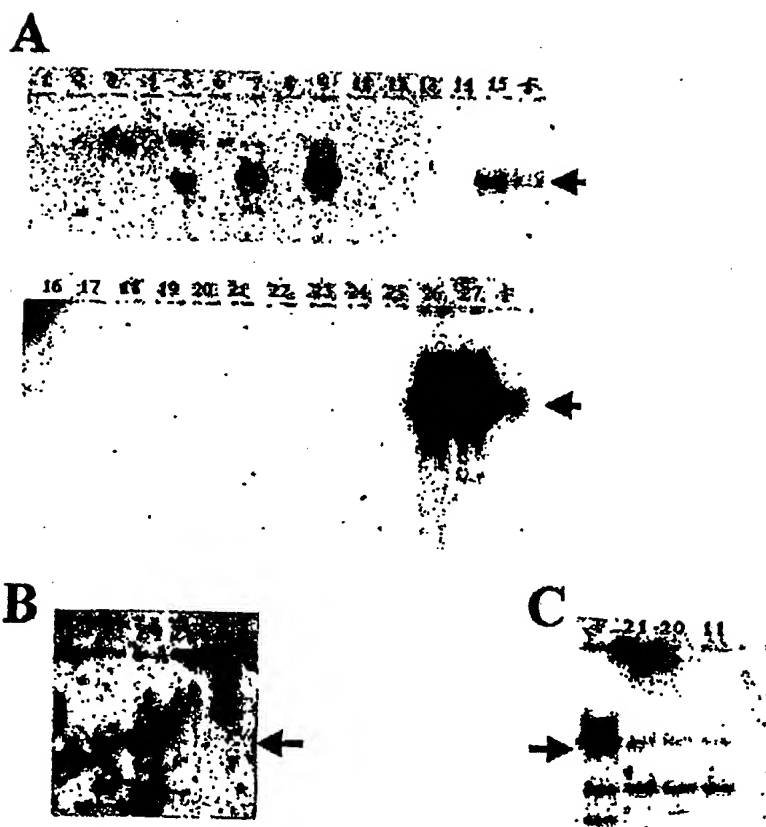


Figure 41

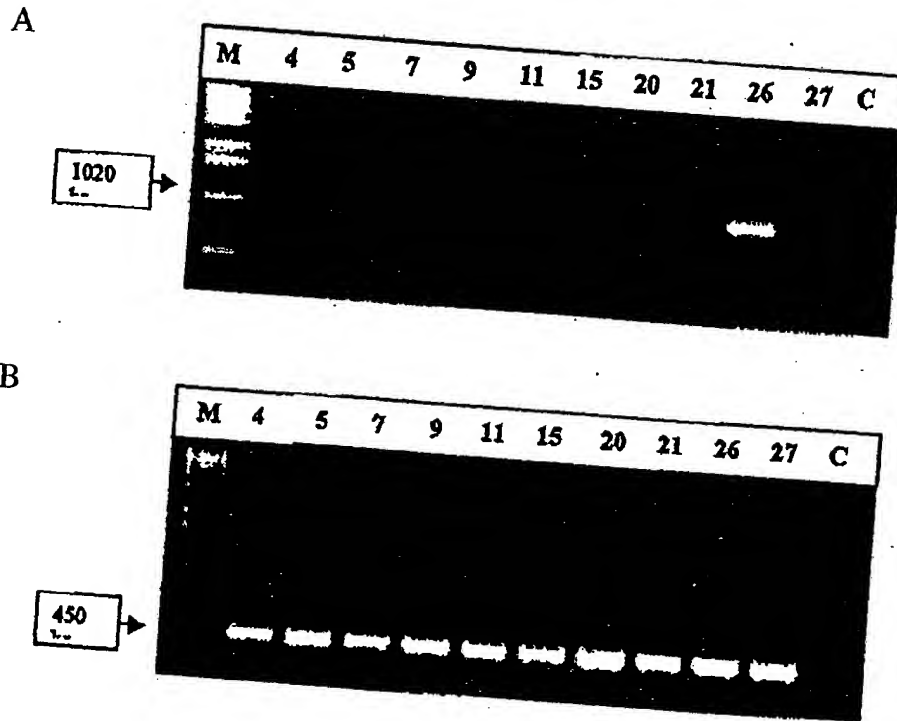
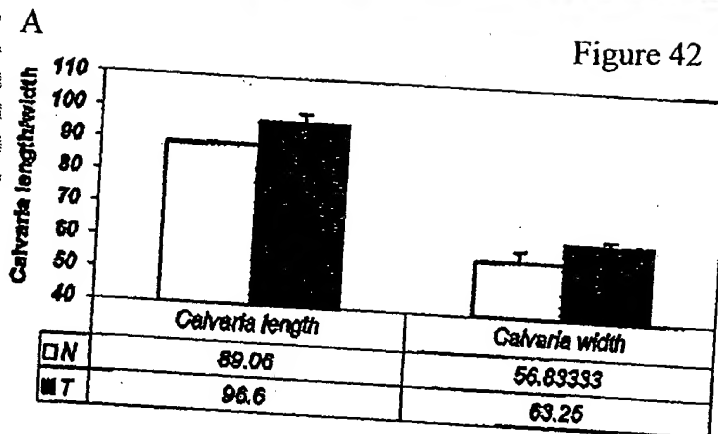


Figure 42



B

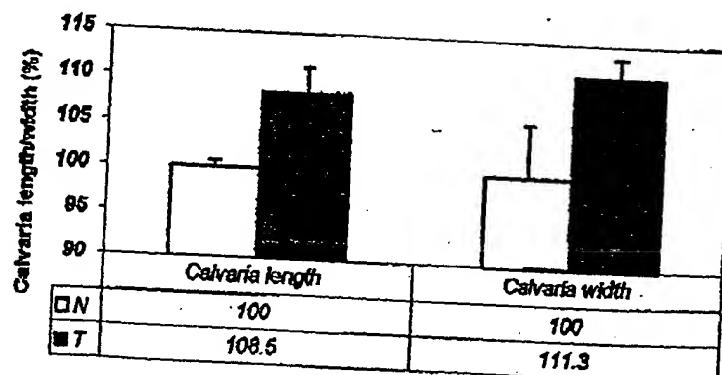


Figure 43

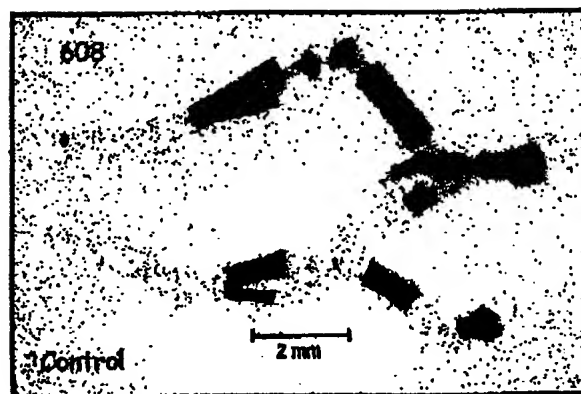


Figure 44

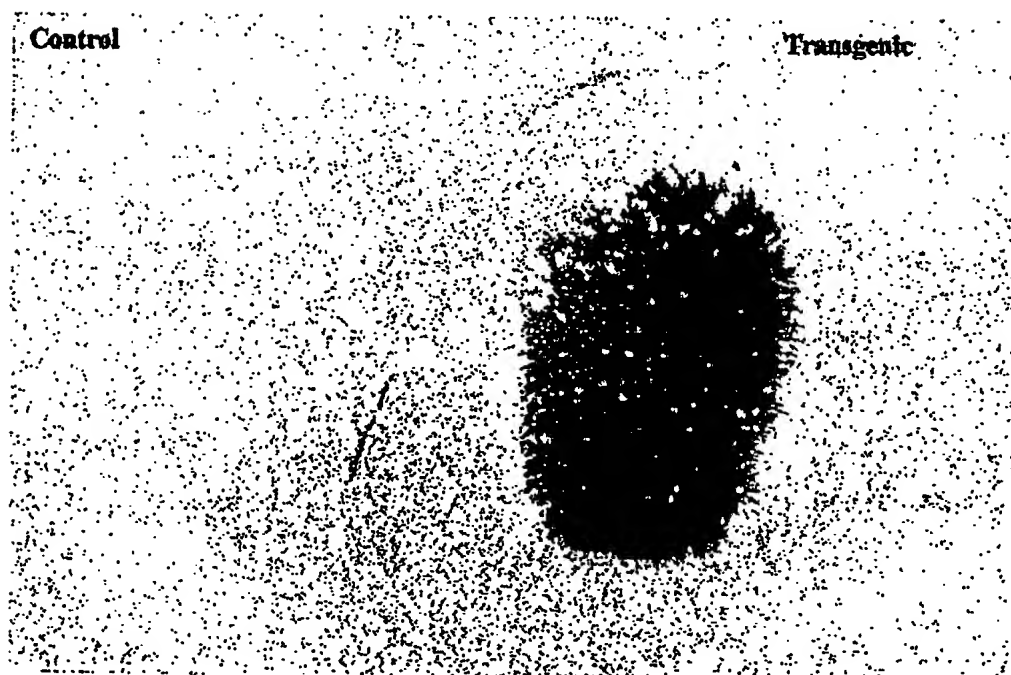
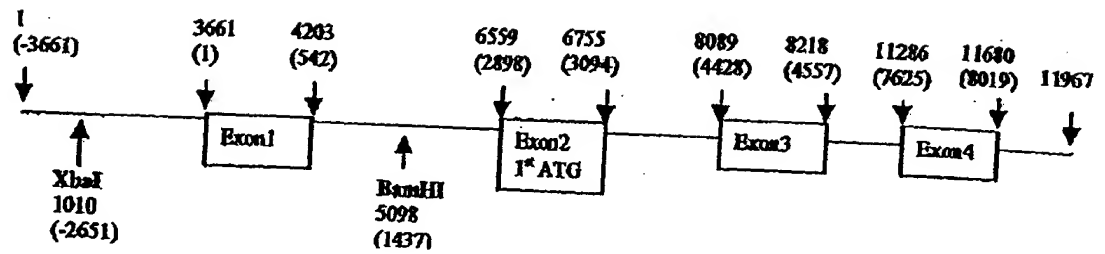


Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46

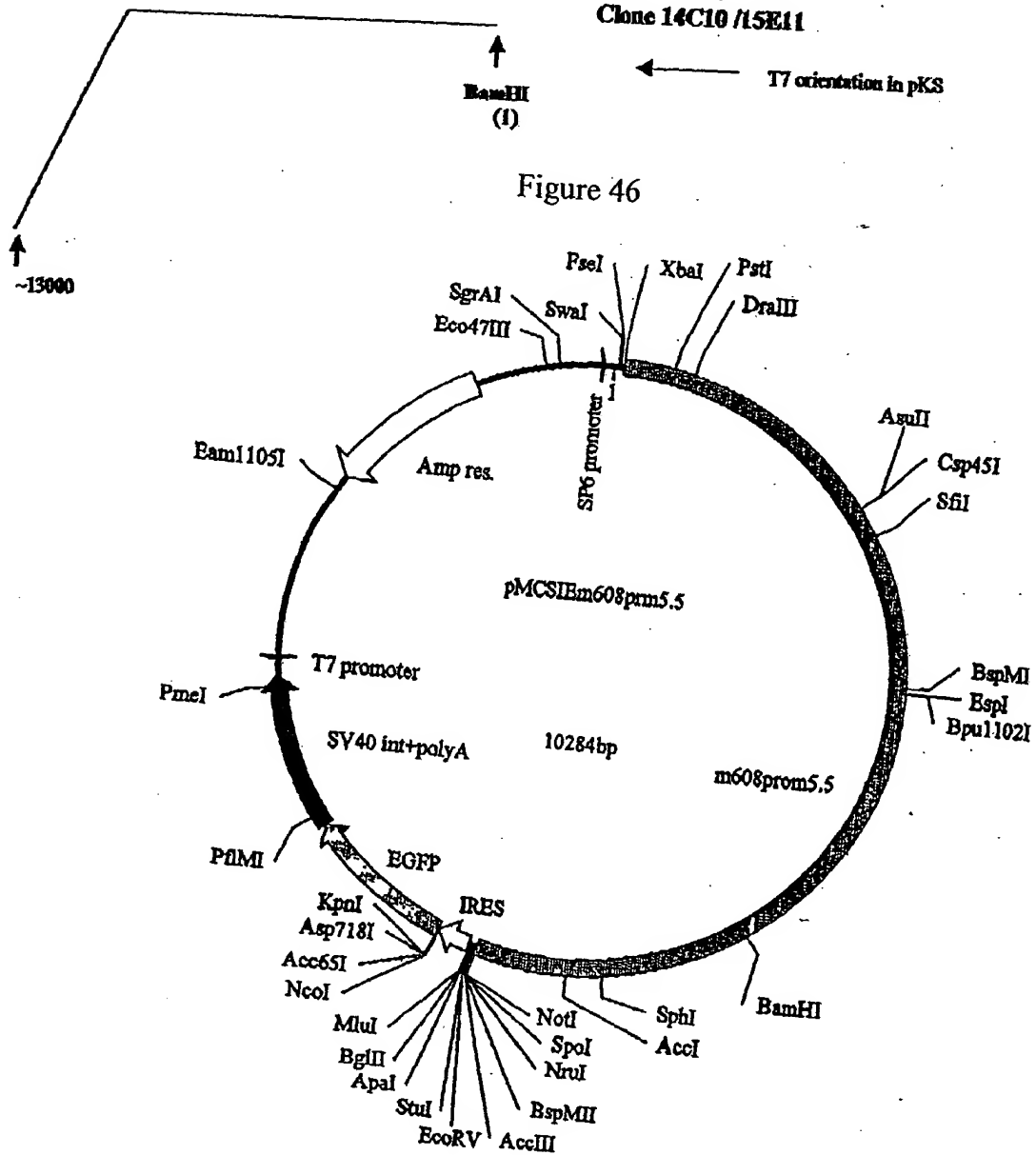


Figure 47

↓ (XbaI)

TCTAGAAGTAAATGATCCTGAGTAGCGATCCTGGGAAAAATACGTACTCTAACACACTGC
 AATCATCTCTCTGTGGTTTGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGTTAGAAATA
 ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCT
 GTAGCTGCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCACATGCCGACAT
 CTTAAGAATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAG
 GGTIAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTACCTTGATCATTAAATGACTAGAA
 ATCACTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAAGCCAGGCCAC
 CTCAGCTTTGGCAGCTCCATTCACTCTGTGGAGCCAGTCACGTGGGTTTGAATCAACTGT
 ACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTTCACAAGCTTTTTTA
 GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG
 TGTTCATGTATGCTCATGGATGCTATTATAATGTGGAAATAAACTGACTTTAAAAA
 AGCTTATGATACTTGTCACAGAGTAAATCTTCCATAAAATATCATCTGCATTTATAAATTAT
 TTTCAATAATCATCAATTAATAAACCCTTTAGAAAATTTGTTAACACAAAGATCCCTAGGCC
 CTGCCCTAGGATGGTCTGTATGGTGGGCTGAGAGATGGAGCTTAAGAACTTACTTGCTCC
 AGGAGCACATCTTCAGAACATCTGCCTCAAAACATTTATCCAAATGCTCATCAAAGGCTC
 ACTCACATGTGCTTCAACCACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGT
 GGAAGCCTGCTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC
 ATAACCCTCACCAACAACAAACAACGACAACAAACCCATAAAATATCAAGGCAAAAT
 GAACTAAGCCATATGCAGAAAAAGTATTATATGTTCTCATTTGTGGGGTGTTTTCTTAAT
 AGTCAAAATATGCAGAAATATAGCAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATA
 CTTGTAGATTAGAGGACACAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAATG
 TATAATATGAGGACTGTATTTAATAATTCTATTAAAGATACACAGCAAAAGAGTGTATCTT
 ACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTTTTATCTTCATG
 TAGCTGATAATTTCATATTGTACACCTCAAACATAGATAACCAACAAAGAGGAAGAGGAT
 AGGTGCTCTCCAGGGGCGGAAGAGTACATTCGAAAGTCAGACACCATTGTGTAGATGTA
 CCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCTATAGG
 TGAACAACATTATGAGCTAACCCAGTACCCCGAGCTCTTGACTCTAGCTGCATATATATC
 AAAAGATGGCCTAATCGCCATCACTGGAAAGAGAGGCCATTGGACTTGCAAACTTTAT
 ATGCCCCAGTACAGGGGAATACAGGGGCCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTG
 GGGGTGGGTGGATATGGGGGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATAC
 CTAATAAAAAATGGAAAAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG
 TTCCTGCTATCTCAACTTGCAAAACAGAGGGGAGTCACAAACCCAGGACAAACACAGT
 GATTGAAGCGTCTTTGAATGTTATTGCTGTTGTTGTTTACCACCATCATTAGCATATATTCAT
 TGTGAAAAAATTACGGGCTCTATGACATGTTTTTTTATTCAAGTATATCAGATGCTGTGAGC
 ATATTTGGCAACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACA
 CACCTTTACCTTCTCCTGGG
 CATCATCTGCTCACTACCCACCCAAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC
 CTATTTTATGCTAGGTTCCCTCTCCCTGTTAGGAGATGGGAGAGGTACGAAAAGGAA
 AGAATTTGTAGCCCTGAGCCAGCCCGGCCACAGAGCCTGCCACCAGACAGGAAAGGCC
 CAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGGGTTGGCGCTGGTGGT
 CCGGGGTGATGGCCCGCCCATTCOCAGAAGCCGAGGCTATAGCTGCGTCACTGCCCCG
 CCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA
 TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAAG
GGTCCAGGAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA
CATCGCTTCAGTGAAGCCAGGCCAAAAGGAGCGGGAAAGGATTATGTAGCCTGGGA
CGCTTTCATAAAACATGATGACGTGTTTGTGCAAAAGCAAGCAATTTGAGGAGAAACGGCT
GGGCGTCCGAAAGAGGAGTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTGAGAAC
TAACCTAATCTATCTATTGAGGGAGAAAGGAGGAGCATTCCAGCAGCAGCAGCAGCA
GCAATCAGATAAAGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAATAACA
GAAGCGCTCCAGTTCTCTGAAGAGTCAGTCCCCAGCTAGTGAAGACTAAGCCTACTAA
GCTTTTGGTCCCGTTGGAAGCAAAAGAGTTCCTTCAATCAGGTGAAGGCTCTCTCAGA

AGATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAAGCAAGACAGAAAGAGCTCAGG
TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTAACGGGAACATCATGTGTA
TTCAATTTACGGCTCTCTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTCTGTTTCC
CAGCTACCTTGTAACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
AOCGAGAGAGTATTCAAAAATATTCAAGTATGTAACAGTGACAGTGTAGGCATAGAAAGTA
ATCATTAGTAAATCTTAATTTGGGTAAACTCATTATAACAGCTCCAGGTTGGGAGGGAT
CACTGAGCCCTTCGCCAGTGCGGGTTAAAGATATTTTCTAACAAGAGAAAGCAGAAATCTTC
CTTGCCCATGCTCCCCATCACTGTGTCAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAG
AGCAGACAGTGTATGCCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAAGTTTACT
GTCTCCCGTCAATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGC
AGACATGCTAGAAAGAAAATTTGACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTAT
TTTGGAAGGTAAAGCCAAATAAATAAACACATATTAATAATACAATGTTACAAAATTT
GATCATATAAAGAAGTACATTCTAATAATGCAATGTGAAAAATATATATAATTTTATCTAT
TTACTGGTGCAAAGTTTTCTAAATTGACATGTACTATTTTATATTTATAAAAAATTTTTT
AAAATGTATATAAAGGTGTAAAAGGCTCTTGGTCAACAAGAGAGTTAAATTTACAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCATTAGTGATAAAGTCCAATG
GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA
CTCTAGGGAATAAAGAGTTCTCTCTCTCCAGATTTCACATAGCATTGTTGTGTCATCTTATC
TTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCTTTATTGGATAA
TTGGGGGATGTTGTTAATTTTTCAGTTCTTATGGAATTCAAAAAACAAAAACAAAC
AAACAAACAAAAACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
AGAAATCTGCACAGAATTTATTGTACATTGTTCAATTATTCACGACAGCCAAGAAATGTGGA
ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAATGTGTACAAC
AGAGTCCCATGTGGCCATAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTAGAATTTTATT
TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
AGCTAGCTGGGTGGGTAGGAGAGAAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACACAGACAGGA
GGCACACGTTCAAACAGGGTGAAATCCAGCACAGAGAAAGGGGAAAGTAGACACAAAGT
TTGCCCACTAACCAAGAAGCCATTTCAGTTGCTGCTGCTGGGAGGGGCGTTCCAGTTT
CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGG
TTTTTGTGCTATTTTTCATTTTCTTACTGTTTTGTTGTTGTTGTTGTTGTTGTTGTTG
GCTGTGTTTTTCATTTGTTTTTGTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
AGCTGGAACGATCTGGAAGAAAGTTGGGAAAGAGAAAAATTTGTATGGAGCATATTTAA
ACAAACAAACAAACAAACAAAGGTTTCAATTTGCCACAAAAGGTGTGAATTAATTAAC
CAGTTACGACTCTTAAGAAAAATTTCCCAATTATCCAGAGTTGCTATGTATGCTGTGC
TAGGACTTTGCTTGAACGGCCATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial

NotI site

(SEQ ID NO:17)

GCCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
AOCCTTGGGGGCTCTCAGAGTCTGAACCAACCAACCATAGAACATTCATGGGCTGTACCCAG
GOCCTCTOCACTCATATGTAAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGOCCTAGTCTGGCCTCAGTGGGAGAGGAAGCAOCTAGCTTTAGCCTT
GTAAAGACTTGAAGTTCGAGGTGTGCGTGGAGGGTATACTCAGGGAGGGCCTCAOCTGC
TAGGAAGAGAAGAGGAGGGGGAAGACTTGGGGGAGGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAAATAAAAAATAATTAAAAAAAAAAAAAAAGAAATGGCTGCTGAGCCC
TACTCTAAAAACCATTGCATCCCCCCCCCAATCATTCACTGACTACGAATTAAAAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTAACTTCTTTGTGAACAAGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTCTTTTCACTTTCCCTGTACATGCATG
TAGCTTAATGATGTTTCAGTGAATTAATAAATAACAATGAAGAATAAAGACAACGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCAOCTGGAAGAAATGTCTTCAGTT
TTTCTATATATGTTTACTCTATOGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACCTTATAAGAAAGTAAAAACATTCATGAAATTGTGAAGGTACTTTTC
CTTACCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAAAACATAGAAGTTGAAATTAAGAAAGAGTAGACAGGA
GAAGTAATAAGGCAAAATAATGAATATTTGGCTTAAATAGTTCCTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAATATCCAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAAATAATAACATTTCTACTTGAATAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

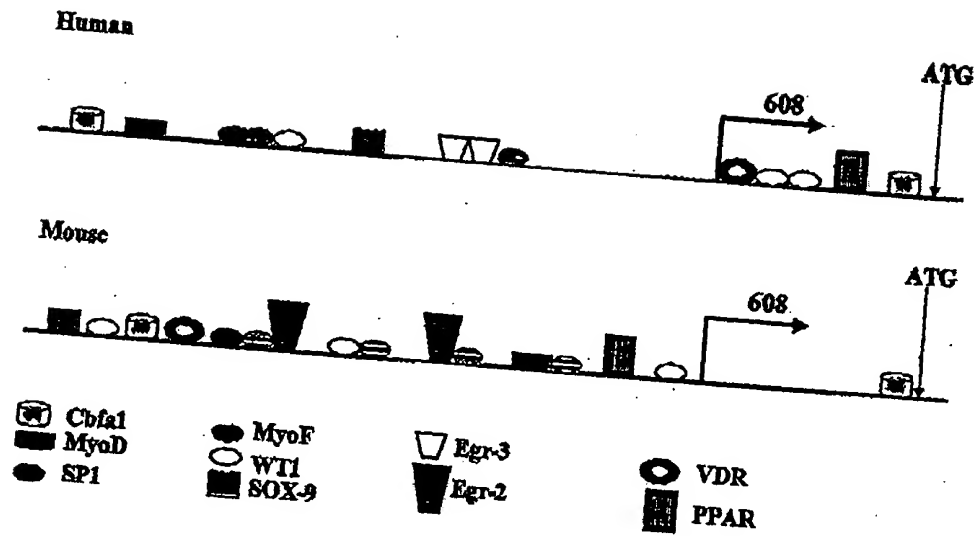


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTTCATAAAATTTT
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAACTGAGTCTTAAGGATTATATTAACCTCCAGTTCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTAACCTTACCAGCTTCTACACAGAGCA
GTTACTATCCTTGCCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAGGG
TTAGTGCTAACTCTGACTGCTTGACTCTTAATTCTCCCTTGATCATTAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GCCACCTCAGCTTTGGCAGCTCCATTCTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTT
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGTAATGTGTTCATGTATGCTCATGGATGCTATTATAAATGTGGGAA
TAAACTGACTTTAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAACCTTTAGAAA
TTTTGTTAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTTCCTTAATAGTCAA
ATATGCAGAAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT
ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAAACCTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAAATGGAAAAAAA
AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACCTTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCCGCCCCAACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTACCCACCCAAGCTTAATCCTTTTCCTCCCTGCAATAGTACCTC
TCCTATTTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTACGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTTCTTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACATCATG
TGGTATTCAATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTAAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTCTATAA
CAGCTCCAGGTGAGGAGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGGCCATGCTCCCCATCACTGTGTACGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCTCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGGAAAGGTAAAGCCA
AATAAATAAAACACATATTAAATAATACAATGTTACAAAAATTGATCATATAAAGA
AGTACATTCTATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATATTTTTAAATG
TATATAAAAGTGTAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTCTCCAGATTTCACTAGCATTGTGTGT
CATATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTCTTTCAGTTCCTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTTGTGCTATTTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAAGGTTTCATTTT
GCCACAAAAGGTGTGAATTAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCTTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTTAAAAAAGCATTTCCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTGTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTCCTTAATTAGTGATTTCATGGG
GCAGGGCCCAGTCCATTGTTTCGTGGTACCATTTCAGGCACTATTAAAAAAAAAAAA
AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCCCTACCCTATTTGAGTTCCTGTCCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTTCCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
AATGCAATTTTATGAGAAAAATTTGAATATTAAGAAAATGCTGGGGAAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTGAGAATCCCAGAGACT
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
TCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAGGAAGG
AGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
GAAAGGGGGAGAAAGAAGAGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTACACACCCGAAGTGGGAAACTTAA
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACAATAAAGTACATAT
GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTTACAGAGACTTTAAACATGAAACTGGAACATATGTCTAGTGTTTGA
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTTCATGTATCCTGCATGAAATTC
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTATTAGTATTTTCATGTTCTTT
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTT
TTTGCTTCTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCCGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TTAAAAAATAAAAAAATATTGATGTGGAATGATCTTAATTTGGGGAAAAGACAATT
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAATG
AGATGGTAGAAAGGAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT
GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATCTAA
AAACAAAACATTTTTTGACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
GGAAGTTGAAGTTCAATTGATAGTCTCATAAATCTTAAATGTATTAACTGCTTTTTA
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCCCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCAATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACCTATCTTGTGTTTGTGTTCTTTTTTT
ATARKACGTATTTTCTCAATTTCAATTTAGAATGATATCCCAAAAGTCCCCCATAACC
TCCCCCCCCTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTGTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Figure 51

MPKRAHWGALSVVLLLLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
LSTIRHLYLAENMVRTLPASMLRNMPLLENLYLQGNPWTCDCEMRWFLWDKSRGIL
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEEEQEQ
EEDGGSQLEKFQLPQWSISLNMDEHGNMVNLVCDIKKPMDEVYKIHLNQTDPPDIDIN
ATVALDFECPMTRENYEKLWKLIAYYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL
YYTGVRQAQILAEPEWVMQPSIDIQLNRRQSTAKKVLLSYTQYSQTISTKDTRQARGRS
WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPDSKFSIL
SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDVTIGKNPGES
VTLPCNALAIPEAHLWILPNRRIINDLANTSHVYMLPNGTSLIPKVQVSDSGYYRCVAV
NQQGADHFTVGITVTKKGSGLPSKRGRPGAKALSRVREDIVEDEGGSGMGDEENTSR
RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTTPFFPAVSPPSASP
VQTVTSAEESSADVPLLGEEEHVLTGISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
SEKTEEITSTEGDLKGTAAPTILISEPYEPSPTLHTLDTVYEKPTHEETATEGWSAADVGSS
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLETKSQPDDEKMKEDTFAHLTPTPTIWNDS
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKGKEMSQTLQGGNM
LEGDPTHSRSESEGEQESKSITLPDSTLGIMSSMSPVKKPAETTGTLLDKDTTITVTTTPR
QKVAPSSTMSTHPSRRRPNGRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGTPRRKHGKRPKNHRYTPSTVSSRA
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDMTTTRKIYSSYPKVQETLP
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTDPLLKELEDVDFTSEFLSSLTVSTPFHQEE
AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS
TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTOHMSGPNE
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFTNTKLSF
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPPLLHTPQTTGSPSTN
LQNPIMVSSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPASKFWSLGEKPQILTKSPQTVSV
TAETDVTVPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTDVRYG
GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVYKSVTDK

[illegible]

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
 TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
 ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
 GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
 ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
 TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCCGATGG AGCTTTAAGA
 GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
 GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
 TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
 ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
 CCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTT AGACTCTCCA
 CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
 AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
 TCCGTGGACC TGCATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
 AATCCAGAGG AATTCTGAAG TGTAAAAAGG ACAAAGCTTA TGAAGGCGGT
 CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
 ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
 TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
 GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAAGTGC CCCAGTGGAG
 CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
 GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
 GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
 TCCAATGACC CGAGAAAAGT ATGAAAAGCT ATGGAAATTG ATAGCATACT
 ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
 CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA
 CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
 AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
 GTGCTACTTT CCTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
 TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
 CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
 TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
 TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
 TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
 GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
 ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
 CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT
 GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
 GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
 GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCAG GTGCAAAGGC TCTTTCAGG GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCTG GAAAAAGAAG CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCCGT TGATTAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTTCTGCTG TTTCTCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTAAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAACATACT TTTGCCTAGA
ACTGTTTCTC TGAAAAGTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
GCTTCAAGTC AGGCAGAAAC CACCACCTT GATCAAGATC ATCTTGAAC
CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
CTGCTGCCCC GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
TGGGGAATCC AGAAACAGAA GCAACCCAG TCAACAATGA AGGAACACAG
CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
TAACTTCCCA GTCACCTCGT CACTGGACCA ACAACCGGA AATAACTACA
TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATT
CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTA CTCCAAA
GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAGCC
TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTACCCG GAGACCCAG
ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
TGAGACAGAC ACTGTGTTCC CTTGTGAGGC AACAGGAAAA CCAAAGCCTT
TCGTTACTTG GACAAAGGT TCCACAGGAG CTCTTATGAC TCCGAATACC
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCC CAAATTTCT
GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTACAT TACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCAATCCCG AGATCTCCTG GAGCCTCCA
GACGGGAGTC TGGTGAATC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACTCTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACCTCTC TTATTAGAA AGCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTC ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCCAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACCTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
AGGGAAGCCA GGTTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCTCTG TCACTTCAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGTACCAT TTTATATGAA AAAAGAAAAA CATTCTTCC
TGGAATCAC TTTTATATA ATGTTTTATA TATATATTTT TTCCTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAATAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTTT TCTCCAACCT CTTCAAATT CAGTCACCAC
TGTTATATTA CTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCTT GTATGCAAAG TTTTGTGTA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Figure 53

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGTGGTCTCCTTTGCTGTGATCTGCCTGGTGGCCACC
CCTGGGGGCAAGGCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTCAGAAAGATACCTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAATATTGAGTTTATAAACCCAGAGGTTTTTATGGGCTCAACTTTCTCCGCTGGTGCACTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCTCAG
CAGTGTCCACTTTTGATGAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCT
TTCCAGTGTGCCAAGCCAACCATTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGT
AGTTCTGCTTTCATCTCTCCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCCATTGCATTCT
ACTGAAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAGATTAC
GGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACCTAGAAAGGAGC
CACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATT
TTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAAATTTCTTGCAG
CTGAACAGAATGCCACCACATTCACTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTA
CCAAGAGCAGAGATGAGGCCAGTGAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGTTA
GAACACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACCTGCCAGGCAAGGAGACCCCAACCCACAC
GTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCCTTATGTGATGAGGATGGACGGATCCTA
ATAGACAAAAGTGGAATAATGGAATCCAGATGGCTGATAGTTTGTACACAGGCGTATATCACTGTATA
AGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTCTGAAGCC
TATCAGGAAAATGGGATTCTACACACAGTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACT
GGTATCCCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGAC
AAGAAAGTTCTAAACAATGGCACATTAGAATATTACAGGTCAACCCGAAAGACCAAGGTTATTATCGC
TGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGGATTTTCCAAGTTTCAGTCAAGATGAAAGGACAA
AGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTT
AAGGAGCCACCAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCA
AGCACAAGTAAGAGGCACAACCTATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGT
TTTAGGGAGAATTAGGGCATTTCCTCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTG
TTGGAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCCACCCCAAGTG
GTCACCCAACCTCCCAACATACCTGGTGAAGAAGACGATTCCCTCAGGCATGCTCGCTCTACATGAGGAA
TTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGACTGCTGACTCCAGAACAATA
TCTGATAGTCCATGACAAACATAAATTATGGCACAGAATTCTCTCTGTTGTGAATTCACAAATACTA
CCACCTGAAGAACCACAGATTTCAAACCTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATA
AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTTCATCCACTGTCTTTCCACTGCTA
CTTGGAGCAACTGAATTTTCAAGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCC
CCAATAACAGTAAGGACTATGATCAAGATGTCAATGTCAAAATGCTTAGTAGCACCACCAACAACTA
TTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCAC
AATCACTTCTATTCTCACACTACTCAAACTAGTACACCTCCACGTTCCCTTCAGATCCACACAGCT
GCTCATTCTCAGTTTCCGATCCCTGAAATAGTACAGTTAACAATCCCGCTGTTTCAGACGCTTTGGGAGG
CAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGA
TACAGCATTTTTCAGGTCAACAACCAGAGGTTCTTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTG
CTCAATGTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTT
CCAAGTGCTGCTCCCATCACCTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCT
CTAGTCCAGAATCCACTATTACTACTTGAGAACAACCCAGTGTAGAGAAAACAACACCCACAATAAAA
TATTTTCAAGACTGAAATTTCCCAAGTACTCCAACCTGGTGCAGTCATGACATATGTCCAACATCCATA
CCCATGGAATAAATCTACAAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGA
GATTCAAGTATTACATCGTCACTTTTCAAGGTGCTATCACCAGGCCCAATGACTATTATAGCCATTACA
AGGTTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTA
AGGAATCAACATAAAGTTAGTTTACAAAAAGCAGAGCTGTGATGCTTCTTAAACATCTCCTGCTTTA
CCACAGAGACAAAGTTCCCTTTCCATTTCACCACTTTTCAACAAGGTGATGCAAAATTCATCTAAT
ACCTTGACTACCGCTCACCACTACGACCAAAACACACAATCCTGGAAGTCTTCAACAAAAGAGGAG
CTTCCCTTCCACCCCTTAACCTATGCTTCTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATC
ATATCAACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTCATCTTATGAAACC
CAAACAGAGAGATCTAGAGCACAACAATACAAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGAC
CCAAACATCTCTCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCA

1000 900 800 700 600 500 400 300 200 100 0

TTTCTACTTATCCCAACAGAAAAACAACACTGGGATTTCAGACACAATCACTTTTCATTTCAGAACTCTT
AATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCT
GAAACAACTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCATTAGACACTCCCAT
CCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCA
GTTACTGACACACGTGGCGACTCCCATTTCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCC
TCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTA
CATTCTACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCTCCCTGG
GCAGAAATACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGC
ATGTTGGCTACTACAGGCCTGTCCGAGGCCACCCTCTTGTTCAGATTGGGATGGACAGAAGAACACA
AAGAAGAGTGACTTTGATAAGAAACCAGTTCAAGAAGCAACAACCTCCAACTCCTTCCCTTTGACTCT
TTGTCTAGGTATATATTTGAAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTACTATTCCAGCT
AATCTCAGATGCCTTTCTTCCCTGTGAGCTGTGGAAATCCCCCTGCCACCAATTTCATTGGACAGAGTT
TCAGGACTTGATTATCTAGAGGAAACCAGAATAGCAGGGTCCAGTTCTCCCCAATGGTACCCTGTCC
ATCCAGAGGGTGGAAATTGAGGACCGCGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGAC
CACCTTCATGTACCTTGTCTGTGGTTTCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATC
ACAGTTCATTCCGGAAGCACTGTGGAACCTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTT
GACGGAACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAAC
CCAGGTGGCCAGGATTCAGTGCTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAA
AGGAGGCAGTCATTGTAGGCACCTGGGGTGAAAGTTTAAAACTGCCCTGTACTGCAAAAGGAACCTCT
CAGCCCAGCGTTTACTGGGTCTCTCTGATGGCACCTGAAGTGAACCATTACAGTTTACCAATTCCAAG
TTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACCTTATGAA
TGCATTGCTACCAGTTCCTACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACC
AGCCCCAGGATAGAAGCTGCATCCCAGAAAAAGGACTGAAGTGAATTTTGGGGCAAAATTACTACTGAAC
TGCTCAGCCATCGGGGAGCCCCAAACCCCAATAATGTGGAGGTTACCATTCCAAGGCTGTGGTGCAGCAG
TGGAGCTGATCCACGTCTACCTTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGT
GTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTG
AAACCTGCCAAAAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCGAAGTA
GATTGCAAAAGCTTCCGGCTCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAAC
AATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACTTTA
TACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGTCTATGCCCAGAACACCCTAGGGAAA
GATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGATAAGGCAGAGTAACAAAACCAAC
AAGAGAATCAAAGCTGGAGACACAGCTGTCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATA
TTTTGGTTGCTGCCTTCCAATGACATGATTTCCCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACA
AACGAACCTGTATTAAAGCCACAGCTGTGAGACATTCACAAAAACCACTTTGACTGCAGAGTGGAAGG
ACACCATCTCTGAAGTCACTGTGGATCATGTGCAGACATAATTTTCTCACAGCCCCATACTATGGAAGC
AGAATCACAGTCCATAAAAAATGGAACCTTGGAATTAGGAATTGAGGCTTTCAGATTACAGCCGACTTT
ATCTGTGTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAATGCTG
AGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTG
AATTGCTCTGTGTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAATGGCACACGATTTTCC
AATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTTTTATCATTCTAAAACAACCTCGG
GAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTA
GAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAGGCATCAGTGGAGAATCT
CTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTAT
GTAGTAGACAGGCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAA
GCAACAGCTTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTGGTCATACACTGATTACT
GTTCCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGG
ACAGGGGCAGCCTTTCAGCTCCAATGTGTGGCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATG
CGTGACCACCTCCCTCTCTCAACGGCAAGTAAAGAGGACACATGGAAGTGAGCAGCTTCACTTACAA
GGTACCCTAGTCATTCAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAGAACC
CTTGGTAGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 54

ORIGIN

1	MKVKGRCITC	LLVSFAVICL	VATPGGKACP	RRCACYMPT	VHCTFRYLTS
51	IPDSIPPVNE	RINLGYNLSV	RLMETDFSG	TKLELLMLHS	NGIHTIPDKT
101	FSDLQALQVL	KMSYNKVRKL	QKDTFYGLRS	LTRLHMDHNN	IEFINPEVFY
151	GLNFLRLVHL	EGNQLTKLHP	DTFVSLSYLQ	IFKISFIKFL	YLSDNFLTSL
201	PQEMVSYPMD	LDSLYLHGNP	WTCDCHLKWL	SDWIEKPDV	IKCKKDRSPS
251	SAQQCPLCMN	PRTSKGKPLA	MVSAAAFQCA	KPTIDSSLKS	KSLTILEDSS
301	SAFISPOGFM	APFGSLTLNM	TDQSGNEANM	VCSIQKPSRT	SPIAFTEEND
351	YIVLNTSFST	FLVCNIDYGH	IQPVWQILAL	YSDSPLILER	SHLLSETPQL
401	YYKYQVAPK	PEDIFTNIEA	DLRADPSWLM	QDQISLQLNR	TATTFSTLQI
451	QYSSDAQITL	PRAEMRPVKH	KWTMISRDN	TKLEHTVLVG	GTVGLNCPGQ
501	GDPTPHVDWL	LADGSKVRAP	YVSEDGRILI	DKSGKLELQM	ADSFDTGVYH
551	CISSNYDDAD	ILTYRITVVE	PLVEAYQENG	IHHTVFIGET	LDLPCHSTGI
601	PDASISWVIP	GNNVLYQSSR	DKKVLNNGTL	RILQVTPKDQ	GYRCVAANP
651	SGVDFLIFQV	SVKMKQGRPL	EHDGETEGSG	LDESNPJIAHL	KEPPGAQLRT
701	SALMEAEVVK	HTSSTSKRHN	YRELTLQRRG	DSTHRRFREN	RRHFPPSARR
751	IDPQHWAALL	EKAKKNAMPD	KRENTTVSPP	PVVTQLPNIP	GEEDSSGML
801	ALHEEFMVPA	TKALNLPART	VTADSRTISD	SPMTNINYG	EFSPVNSQI
851	LPPEPTDFK	LSTAIKTAM	SKNINPTMSS	QIQGTTNQHS	STVFPLLLGA
901	TEFQSDQMG	RGREHFQSRP	PITVRTMIKD	VNVKMLSSTT	NKLLLESVNT
951	TNSHQTSVRE	VSEPRHNHFY	SHTTQILSTS	TFPSDPHTAA	HSQFPPIRNS
1001	TVNIPLFRFR	GRQRKIGGRG	RIISPYRTPV	LRRHYSIFR	STTRGSSEKS
1051	TTAFSATVLN	VTCLSCLPRE	RLTTATAALS	FPSAAPITFP	KADIARVPSE
1101	ESTTLVQNPL	LLLENKPSVE	KTTPTIKYFR	TEISQVTPGT	AVMTYAPTSI
1151	PMEKTHKVNA	SYPRVSSTNE	AKRDSVITSS	LSGAITKPPM	TIIAITRFSR
1201	RKIPWQQNFV	NNHNPKGRLR	NQHKVSLQKS	TAVMLPKTSP	ALPQRQSSPF
1251	HFTTLSTSV	QIPSNLTLLA	HHTTTKTHNP	GSLPTKKELP	FPPLNPLPS
1301	IISKDSSTKS	IISTQTAIPA	TTPTFPASVI	TYETQTERS	AQTIQREQEP
1351	QKKNRDTPNI	SPDQSSGFTT	PTAMTPPALA	FTHSPPENTT	GISSTISFHS
1401	RTLNLTDVIE	ELAQASTQTL	KSTIASSETTL	SSKSHQSTTT	RKASLDTPIP
1451	PFLSSSATLM	PVPISPPFTQ	RAVTDTRGDS	HFRLMTNTVV	KLHESSRHNL
1501	QMPSSQLEPL	TSSTSLLHS	TPMPALTTVK	SQNSKLTSP	WAEYQFWHKP
1551	YSDIAEKGGK	PEVSMLATG	LSEATTLVSD	WDGQKNTKKS	DFDKKPVQEA
1601	TTSKLLPFDS	LSRYIFEKPR	IVGGKAASFT	IPANSDAFLP	CEAVGNPLPT
1651	IHWTRVSGLD	LSRGNQNSRV	QVLPNGTSLI	QRVEIQDRGQ	YLCASNLFG
1701	TDHLHVTLSV	VSYPRIER	RTKEITVHSG	STVELKCRAE	GRPSPTVTWI
1751	LANQTVVSES	SQGSRQAVVT	VDGTLVLHNL	SIYDRGFYKC	VASNPGGQDS
1801	LLVKIQVIAA	PPVILEQRRQ	VIVGTWGESL	KLPCTAKGTP	QPSVYWVLS
1851	GTEVKPLQFT	NSKFLFSNG	TLYIRNLASS	DRGTYECIAT	SSTGSERRVV
1901	MLTMEERVTS	PRIEASQKR	TEVNFGDKLL	LNCSATGEPK	PQIMWRLPSK
1951	AVVDQWSWIH	VYPNGSLFIG	SVTEKDSGVY	LCVARNKMGD	DLILMHVSLR
2001	LKPAKIDHKQ	YFRKQVLHGK	DFQVDCASG	SPVPEISWSL	PDGTMINNAM
2051	QADDSGHRTR	RYTLFNNGTL	YFNKVGVAEE	GDYTCYAQNT	LGKDEMKVHL
2101	TVITAAPRIR	QSNKTNKRIK	AGDTAVLDCE	VTGDPKPKIF	WLLPSNDMIS
2151	FSIDRYTFHA	NGSLTINKVK	LLDSGEYVCV	ARNPSGDDTK	MYKLDVVS
2201	PLINGLYTNR	TVIKATAVRH	SKKHFDCAE	GTPSPEVMWI	MPDNIFLTAP
2251	YYSRITVHK	NGTLEIRNVR	LSDSADFCV	ARNEGGSVL	VVQLEVLEML
2301	RRPTFRNPFN	EKIVAQLGKS	TALNCSVDGN	PPPEIIWILP	NGTRFSNGPQ
2351	SYQYLIASNG	SFIISKTTRE	DAGKYRCAAR	NKVGYIEKLV	ILEIGQKPMI
2401	LTYAPGTVKG	ISGESLSLHC	VSDGIPKPN	KWTMPSGYVV	DRPQINGKYI
2451	LHDNGTLVIK	EATAYDRGNY	ICKAQNSVGH	TLITVPVMIV	AYPPRITNRP
2501	PRSIVTRTGA	AFQLHCVALG	VPKPEITWEM	PDHSLSTAS	KERTHGSEQL
2551	HLQGTLLVIQ	PQTSDSGIYK	CTAKNPLGSD	YAATYIQVI*	

Figure 55

OCP rat amino acid sequence

>608-663Nterm Rat Protein (663 aa)

MQVRGREVSGLLISLTAVCLVVTTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVLSYLQ
IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
SASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
NLVCSVDYNHIQPVWQLLALYSDSLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNN
PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQM
ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQ

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663

Figure 56

pCM-H-608-663-N-term

